

Universität Duisburg-Essen
Fakultät für Bildungswissenschaften
Lehrstuhl für Lehr-Lernpsychologie

Geschlechtsunterschiede in mathematischen Fähigkeiten und akademischen Selbstkonzepten bei Kindern im Vor- und frühen Grundschulalter

Dissertation zur Erlangung des Grades Dr. phil.

vorgelegt von Katleen Sahr

geboren am 13.07.1982 in Hagenow

Erstgutachter: Prof. Dr. Detlev Leutner, Universität Duisburg-Essen

Zweitgutachter: Prof. Dr. Susanne Buch, Bergische Universität Wuppertal

Tag der mündlichen Prüfung: 12. Oktober 2012

Inhaltsverzeichnis

Abbildungsverzeichnis.....	4
Tabellenverzeichnis.....	5
Danksagung.....	6
1. Allgemeine Einführung.....	7
1.1 Entwicklung früher mathematischer Fähigkeiten.....	8
1.1.1 Geschlechtsunterschiede in mathematischen Fähigkeiten.....	12
1.1.2 Ursachen von Geschlechtsunterschieden in mathematischen Fähigkeiten.....	15
1.2 Akademisches Selbstkonzept.....	17
1.2.1 Akademisches Selbstkonzept im Vor- und Grundschulalter.....	20
1.2.2 Erfassung des akademischen Selbstkonzepts im Vor- und Grundschulalter.....	22
1.2.3 Geschlechtsunterschiede beim akademischen Selbstkonzept im Vor- und Grundschulalter.....	23
1.3 Forschungsfragen.....	26
2. Studie I - Girls and Math - a never-ending story!?	29
2.1 Introduction.....	29
2.2 Theoretical background.....	31
2.3 Method.....	39
2.4 Results.....	40
2.5 Discussion.....	44
3. Studie II - Erfassung des akademischen Selbstkonzepts bei Erstklässlern durch eine adaptiv gestaltete Mitschülervergleichsmethode.....	47
3.1 Forschungsstand zum akademischen Selbstkonzept.....	48
3.2 Methode.....	55
3.3 Ergebnisse.....	62

3.4 Diskussion.....	66
4. Studie III - Unterschiede im Selbstkonzept zwischen Mädchen und Jungen in der 1. Klasse.....	69
4.1 Einleitung.....	70
4.2 Methode.....	80
4.3 Ergebnisse.....	81
4.4 Diskussion.....	85
5. Zusammenfassende Diskussion.....	88
5.1 Zentrale Ergebnisse.....	88
5.2 Theoretischer Nutzen.....	92
5.3 Praktischer Nutzen.....	96
5.4 Zusammenfassung und Ausblick.....	99
6. Literatur.....	103
7. Anhang.....	122
A Selbstkonzepterfassung.....	122
A.1 Instruktion.....	122
A.2 Smileys zur Selbstkonzepterfassung.....	123

Abbildungsverzeichnis

Abbildung 1.1:	Struktur des multidimensionalen und hierarchischen Selbstkonzepts von Shavelson et al. (1976); nach Möller und Trautwein (2009, S. 182).....	18
Figure 2.1:	Math achievement (total scores) of boys and girls from kindergarten to 2 nd grade, * $p < .05$	41
Figure 2.2:	Verbal achievement of boys and girls from kindergarten to 2 nd grade, * $p < .05$	42
Figure 2.3:	Residuals of math achievement (controlled for verbal ability) for boys and girls from kindergarten to 2 nd grade, ** $p < .01$	43
Abbildung 3.1:	Beispielhafte Darstellung der Rangfolgen innerhalb einer Klasse und der adaptiven Auswahl der Vergleichsmitschüler pro Kind.....	58
Abbildung 3.2:	Beispielhafte Darstellung der Antworten der Kinder (0 – Kind schätzt Mitschüler als besser ein; 1 – Kind schätzt sich als besser ein) sowie der sich daraus neu ergebenden Rangzuordnungen.....	61

Tabellenverzeichnis

Table 2.1:	Mean score and gender differences in student performance on the PISA math scale (adapted from OECD, 2004, 2009, 2010; Stanat & Kunter, 2001).....	33
Table 2.2:	Number of boys and girls, math ability, verbal ability and residuals (math ability controlled for verbal ability): Means and standard deviations.....	40
Tabelle 3.1:	Multitrait-Multimethod-Matrix der Selbstkonzeptmessung (Rechnen, Lesen und Schreiben als Traits; Lehrereinschätzung, Fragebogen und Mitschülervergleich als Methods; ergänzt um Korrelationen mit einem Mathematiktest).....	65
Tabelle 4.1:	Mittelwerte und Standardabweichungen für Lehrereinschätzung, Selbstkonzept durch Fragebogenitems, Selbstkonzept durch Mitschülervergleich und Mathematiktest.....	83
Tabelle 4.2:	Korrelationen (Rangkorrelationskoeffizienten Rho nach Spearman) zwischen Lehrereinschätzung, Leistung im Mathematiktest und dem Selbstkonzept in Mathematik (Fragebogen & Mitschülervergleich) getrennt nach Mädchen und Jungen.....	84

Danksagung

Zunächst möchte ich Prof. Dr. Dr. Detlev Leutner, Dr. Maria Opfermann und Prof. Dr. Annemarie Fritz-Stratmann für die freundliche Aufnahme in das ANAC-Projekt danken. Diese Dissertation wäre wahrscheinlich ohne die Projektarbeit weder so noch in ähnlicher Form entstanden. Aus dem gleichen Grund danke ich allen Kolleginnen und Kollegen für ihre fachliche, methodische und auch praktische Hilfe während der Projektphase sowie der Erstellung der Dissertation.

Ein besonderer Dank geht dabei an:

Dominique, ich werde unsere stets positive Zusammenarbeit in bester Erinnerung behalten. Vielen Dank für die Erklärung der Ruhrgebietstraditionen (...und es schneit doch!) sowie der Uni Duisburg-Essen-Abläufe (...t steht für tannengrün), für innovative Süßigkeitenideen (Es sind noch Hello-Kitty-Bonbons da) und deine bis heute andauernde Bereitschaft, mir fachliche und auch moralische Unterstützung zukommen zu lassen!

Maria, vielen Dank für deine unermüdliche Unterstützung sowohl beim wissenschaftlichen Arbeiten als auch bei (fußballerischen) Niederlagen. Ich freue mich sehr, weiterhin mit dir zusammenarbeiten und verschiedenste Freizeitaktivitäten erleben zu dürfen!

Theresa, ich kann mich glücklich schätzen, auch in dir eine tolle Bürokollegin bekommen zu haben. Sich in kritischen Phasen gegenseitig aufzubauen, ist viel wert, weshalb ich dich sicherlich nicht allzu lang allein im Büro lassen werde!

Außerdem danke ich Prof. Dr. Dr. Detlev Leutner für die ertragreiche Betreuung als Doktorvater und Prof. Dr. Susanne Buch für die spontane Bereitschaft, diese Arbeit als Zweitgutachterin zu bewerten.

Nicht zuletzt möchte ich mich bei meiner Familie und meinen Freunden für ihr stetes Interesse an meiner Arbeit und meinem Wohlbefinden bedanken. Peter, ich danke dir für das Zuhören, das Nachfragen, dein Verständnis und deine bedingungslose Unterstützung!

Durch euch wurde und werde ich immer wieder auf meinem Weg bestärkt!

1. Allgemeine Einführung

Mathematische Fähigkeiten sind entscheidend für die schulische und berufliche Entwicklung. Ein Ziel der schulischen Ausbildung ist u.a., dass Schülerinnen und Schüler¹ die gleichen mathematischen Fähigkeiten erwerben und zeigen können. Dass dies nicht selbstverständlich ist, zeigten zahlreiche nationale sowie internationale Studien (z.B. Else-Quest, Hyde, & Linn, 2010; Hosenfeld, Köller, & Baumert, 1999; Hyde, Fennema, & Lamon, 1990; Mullis et al., 1997; OECD, 2004; Winkelmann, van den Heuvel-Panhuizen, & Robitzsch, 2008). Allerdings geht es dabei nicht nur um die Chancengleichheit für Mädchen und Jungen, sondern auch um den schon heutzutage bestehenden Fachkräftemangel in den MINT (Mathematik, Informatik, Naturwissenschaft, Technik)-Berufen (Institut der deutschen Wirtschaft Köln, 2008). Fend (1991) hat in einer Längsschnittstudie zeigen können, dass die Berufswahl von Frauen und Männern geschlechtsstereotypisch geprägt ist. So wählten Männer bevorzugt Berufe im MINT-Bereich (Elektro- und Computertechnik, Metallverarbeitung, Baugewerbe), während Frauen eher Berufe im Gesundheitswesen sowie in Bildung und Erziehung favorisierten. Dass die Berufswahl der jungen Frauen selten auf mathematisch-technische Bereiche fällt (Halpern et al., 2007; Hannover, 1991), scheint allerdings schon in der schulischen Laufbahn begründet zu sein, denn Mädchen entscheiden sich schon in der Sekundarstufe in den MINT-Fächern lieber für den Grund- als den Leistungskurs, während Jungen mehrheitlich den Leistungskurs in diesen Bereichen wählen (Hannover, 1991; Köller, Daniels, Schnabel & Baumert, 2000; Sherman & Fennema, 1977). Zudem konnte nachgewiesen werden, dass es einen hohen Zusammenhang zwischen den gewählten Leistungskursen in der Schule und dem

¹ Im Folgenden wird zur besseren Lesbarkeit der Begriff „Schüler“ verwendet.

späterem Studienfach gibt (Heine, Kerst und Sommer, 2007). Die Ursache dieser Geschlechtsunterschiede tritt also bereits während der Schulzeit oder sogar noch früher in Erscheinung. Der immer wieder bestätigte Zusammenhang zwischen akademischen Leistungen und den entsprechenden Selbstkonzepten, also den persönlichen Einschätzungen eigener Leistungen, könnte eventuell schon früh auftretende Geschlechtsunterschiede erklären (Calsyn & Kenny, 1977; Dickhäuser, 2006; Helmke, 1992). Je früher sich Mädchen und Jungen in ihren Leistungen bzw. Fähigkeiten oder in ihren Einstellungen dazu unterscheiden, desto ausgeprägter wird die Problematik beim Übergang von Schule in Studium und Beruf. Deswegen ist es notwendig zu klären, ab wann Geschlechtsunterschiede in mathematischen Leistungen und dem zugehörigen Selbstkonzept auftreten.

In den folgenden Abschnitten dieses Kapitels werden zunächst die mathematische Entwicklung bis zum Grundschulalter und die bisherigen Erkenntnisse zu Geschlechtsunterschieden in mathematischen Leistungen dargestellt. Daraufhin folgt eine Betrachtung des akademischen Selbstkonzepts und entsprechender Befunde zu Geschlechtsunterschieden. Aus den sich daraus ergebenden Forschungsdesiderata werden anschließend die zentralen Forschungsfragen entwickelt.

1.1 Entwicklung früher mathematischer Fähigkeiten

Bereits innerhalb des ersten Lebensjahres konnten in zahlreichen Untersuchungen grundlegende mathematische Fähigkeiten bei Babys nachgewiesen werden (Baillargeon, 1993; Baillargeon & DeVos, 1991; Barth, Kanwisher & Spelke, 2003; Feigenson & Carey, 2003; Feigenson, Carey & Hauser, 2002; Hespos & Baillargeon, 2001; Kellman, 1993; Spelke, Kestenbaum, Simons & Wein, 1995; van

Oeffelen & Vos, 1982; Wynn, 1992, 1995; Xu & Arriaga, 2007; Xu & Spelke, 2000). Als eine elementare Voraussetzung für die mathematische Entwicklung erwies sich dabei, dass Säuglinge schon kurz nach der Geburt verschiedenartige Stimuli wahrnehmen und unterscheiden können (Baillargeon & DeVos, 1991; Kellman, 1993; Wynn, 1992). Im Alter von 2 ½ Monaten konnte bereits das Verständnis für das Prinzip der Objektpermanenz (Baillargeon, 1993; Hespos & Baillargeon, 2001) und der diskreten Objektrepräsentation (Spelke et al., 1995; Wynn, 1992) nachgewiesen werden. Dies bedeutet, dass Säuglinge eine Vorstellung von Objekten über eine kurze Zeit aufrecht erhalten und dass sie dargebotene Objekte als diskret wahrnehmen können.

Die numerischen Fähigkeiten von Säuglingen basieren auf zwei Kernsystemen (Feigenson, Dehaene & Spelke, 2004). Das System der ungefähren Größenrepräsentation ermöglicht es Kindern ab einem Alter von fünf Monaten, zwei Mengen, deren Mächtigkeit mindestens ein Verhältnis von 1:2 aufweist, zu unterscheiden (Barth et al., 2003; van Oeffelen & Vos, 1982; Xu & Arriaga, 2007; Xu & Spelke, 2000). Das zweite Kernsystem erlaubt die präzise Repräsentation von kleinen Mengen. Dies zeigt sich darin, dass Kinder ab sechs Monaten nicht nur bis zu drei Objekte präzise erfassen, sondern auch einfache Additions- und Subtraktionsaufgaben im gleichen Zahlenraum durchführen können (Feigenson & Carey, 2003; Feigenson et al., 2002; Wynn, 1992, 1995). Somit sind die beiden Kernsysteme, welche auch im Erwachsenenalter noch nachweisbar sind, als angeborene Grundlage für die weiteren zu erwerbenden mathematischen Konzepte anzusehen. Mit Beginn des Spracherwerbs lernen Kinder die Zahlwortreihe, was allerdings zunächst ohne ein ordinales oder kardinales Verständnis der Zahlen geschieht (Brannon & Van de Walle, 2001; Le Corre, Van de Walle, Brannon &

Carey, 2006). Erst ab einem Alter von ungefähr 3 ½ Jahren verstehen Kinder die kardinale Bedeutung der Zahlworte (Brannon & Van de Walle, 2001). Fünfjährige lösen Additions- und Subtraktionsaufgaben so sicher, dass sie mit größeren Zahlen und verdeckten Teilmengen operieren können (Barth et al., 2006). Ein wichtiger Schritt in der mathematischen Entwicklung ist das Teil-Teil-Ganzes-Verständnis (Piaget, 1965; Steffe, 1992). Die Erkenntnis, dass das letzte Zahlwort die gesamte gezählte Menge enthält und dass die Vorgängerzahlen in der letzten Zahl beinhaltet sind, scheint zwischen drei und vier Jahren zu beginnen (Hunting, 2003). Dies ermöglicht Kindern das Operieren mit zwei Teilmengen und einer Gesamtmenge, wobei im Vorschulalter lediglich die beiden Teilmengen sicher zu einer Gesamtmenge verbunden werden können und später auch Aufgaben gelöst werden können, in denen eine Teil- oder Ausgangsmenge bestimmt werden muss (Stern, 1992). Bei Kindern ab ca. 7 ½ Jahren konnte der relationale Zahlbegriff, welcher das Verständnis beinhaltet, dass der Abstand zwischen aufeinanderfolgenden Zahlen immer gleich ist, nachgewiesen werden (Riley, Greeno & Heller, 1983; Stern, 1992).

Basierend auf diesen Erkenntnissen, fassten Fritz und Ricken (2008) die einzeln untersuchten Konzepte zu einem Modell der Entwicklung mathematischer Fähigkeiten zusammen. Grundannahme des Modells ist, dass die für den Erwerb erforderlichen zentralen Konzepte von den Kindern nacheinander und sukzessiv aufeinander aufbauend verstanden werden. Die fünf Niveaus werden im Normalfall im Alter zwischen vier und acht Jahren erworben und ermöglichen somit eine beschulungsunabhängige Fähigkeitseinschätzung. Das Modell lässt sich wie folgt zusammenfassen:

- Niveau I - Zählzahl: Die Kinder können nicht nur die Zahlwortreihe aufsagen, sondern auch kleine Mengen aus- und abzählen. Ist das sichere

Aus- und Abzählen auf vier Objekte anwendbar, so gilt dieses Konzept als erworben.

- Niveau II – Repräsentation eines mentalen Zahlenstrahls: Die Kinder erwerben eine Vorstellung über den Zahlenstrahl, so dass sie jeweils Vorgänger und Nachfolger einer Zahl benennen können. Zusätzlich werden einfache Additions- und Subtraktionsaufgaben nicht mehr nur handelnd gelöst.
- Niveau III – Kardinalität und Zerlegbarkeit: Die Kinder entwickeln ein Verständnis für die Verbindung zwischen Zahlwort und Menge. Ihnen wird bewusst, dass das letzte Wort beim Auszählen für die Gesamtmenge aller abgezählten Objekte steht. Außerdem können Vergleiche zwischen Zahlen nicht mehr nur ordinal auf dem Zahlenstrahl, sondern auch anhand der Menge der Objekte, die die Zahlen beinhalten, durchgeführt werden. Die Kinder verstehen zusätzlich, dass eine Gesamtmenge aus zwei Teilmengen besteht, so dass z.B. Additionsaufgaben ausgehend von der einen Teilmenge gelöst werden.
- Niveau IV - Enthaltensein: Das Teil-Teil-Ganze-Konzept wird präzisiert, indem Kinder ein Verständnis dafür entwickeln, dass Teilmengen aus unterschiedlichen Einheiten oder Klassen bestehen können. Sie verstehen, dass Teilmengen in der Gesamtmenge enthalten sind. Somit wird bei der Subtraktion eine Teilmenge aus der Gesamtmenge herausgelöst und die zweite Teilmenge entsteht.
- Niveau V - Relationalität: Die Kinder erkennen, dass zwischen benachbarten Zahlen der Zahlwortreihe kongruente Intervalle bestehen. Mit dem Verständnis, dass Vorgänger und Nachfolger sich immer um 1

unterscheiden, können auch Differenzen zwischen Zahlen unabhängig vom Nullpunkt bestimmt werden.

Auf Grundlage des Modells wurden bisher der MARKO-D Test (Ricken, Fritz & Balzer, 2011) zur Erfassung der mathematischen Fähigkeiten im Vorschulalter und ein Basistraining für das Kindergartenalter veröffentlicht (Gerlach & Fritz, 2011).

1.1.1 Geschlechtsunterschiede in mathematischen Fähigkeiten

Bezüglich der Geschlechtsunterschiede in mathematischen Fähigkeiten zeigten sich widersprüchliche Befunde. Hyde et al. (1990) wiesen in einer Metaanalyse nach, dass sich der Vorteil der Jungen in mathematischen Fähigkeiten zwischen 1967 und 1987 verringert hat ($d = .31$ auf $d = .14$). Aktuelle Studien konnten diese Tendenz bestätigen, indem sie nur sehr geringe Effekte ($d < 0.15$ bzw. $d < 0.06$) für Geschlechtsunterschiede in Mathematik fanden (Else-Quest et al., 2010; Hyde, Lindberg, Linn, Ellis & Williams, 2008). Ein weiterer Hinweis darauf, dass auf den ersten Blick eine Diskussion über Geschlechtsunterschiede in Mathematik in Bezug auf die benachteiligten Mädchen überflüssig erscheint, sind Vergleiche der Schulnoten. Es zeigte sich in mehreren Studien, dass Mädchen spätestens in der Sekundarstufe zu den Jungen aufholten bzw. sogar bessere Noten als sie erhielten (Kenney-Benson, Pomerantz, Ryan, & Patrick, 2006; Steinmayr & Spinath, 2008; Van Houtte, 2004). Der Rückgang der Unterschiede in den mathematischen Leistungen wurde allerdings hauptsächlich anhand von Studien mit Personen aus Nordamerika belegt (Hyde et al., 1990; 2008; Kenney-Benson et al., 2006). Außerdem zeigte sich bei den Analysen der PISA-Daten zwar ein wenig bedeutsamer Geschlechtseffekt als Mittelwert über alle teilnehmenden Staaten, aber

bei Betrachtung der Werte einzelner Staaten sind große Unterschiede sichtbar (siehe Tabelle 2.1). Während z.B. isländische Mädchen bei PISA 2003 15 Punkte besser abschnitten als die Jungen ($d = -0.17$), erreichten koreanische Jungen 23 Punkte mehr als die Mädchen ($d = 0.26$; Zimmer, Burba & Rost, 2004). Außerdem zeigte sich sowohl bei PISA als auch bei TIMSS, dass Jungen auf den oberen Kompetenzniveaus dominierten (OECD, 2004; Penner, 2008). In Deutschland kann also von einer Verringerung der Geschlechtsunterschiede insofern keine Rede sein, da der Vorteil der Jungen zwar von PISA 2000 mit 15 Punkten bis PISA 2003 mit 9 Punkten gesunken, allerdings bei PISA 2006 wieder auf 20 Punkte angestiegen ist (OECD, 2004, 2009; Stanat & Kunter, 2001). Außerdem zeigte Klieme (1997) im Gegensatz zu Hyde et al. (1990) in einer Metaanalyse von Studien zu Geschlechtsunterschieden aus deutschsprachigen Ländern, dass die Vorteile der Männer im Altersbereich von 16-30 Jahren größer sind als in nordamerikanischen Studien ($0.49 < d < 0.55$) und dass die Unterschiede sich über die Jahre nicht verringert haben. Diese Befunde über Geschlechtsunterschiede in Mathematik, die hauptsächlich Schüler ab der Sekundarstufe betreffen, legen den Schluss nahe, dass sich Mädchen und Jungen auch schon in früheren Jahren hinsichtlich ihrer mathematischen Fähigkeiten unterscheiden könnten.

In den in Abschnitt 1.1 beschriebenen Untersuchungen einzelner Konzepte der frühen mathematischen Entwicklung wurden weder für Mädchen noch für Jungen Aussagen über Fähigkeitsvorteile gemacht, was vermuten lässt, dass sich in dieser sehr jungen Altersgruppe keine Geschlechtsunterschiede bezüglich mathematischer Fähigkeiten gezeigt haben (vgl. Spelke, 2005). Ab der Vor- und Grundschulzeit zeigten Untersuchungen, dass Mädchen und Jungen z.B. bei Bearbeitung komplexer mathematischer Aufgaben die gleichen Leistungen

erbringen (Hyde et al., 1990; Lachance & Mazzocco, 2006). Im Gegensatz dazu finden sich Studien, in denen die Jungen im Gesamtergebnis besser waren, in denen aber abhängig vom Aufgabentyp oder Inhalt der Aufgaben die Mädchen entweder zu den Jungen aufschließen oder sogar besser ab schneiden konnten (van den Heuvel-Panhuizen, 2004; Winkelmann et al., 2008; Zohar & Gershikov, 2008). Auch Tiedemann und Faber (1994) konnten in Deutschland vom Vorschulalter bis zum Ende der Grundschulzeit keine Unterschiede in mathematischen Fähigkeiten zwischen Mädchen und Jungen nachweisen.

Weinhold Zulauf, Schweizer und von Aster (2003) fanden dagegen in ihrer Untersuchung, dass Jungen in der Schweiz bereits ab dem letzten Kindergartenjahr die mathematischen Fähigkeiten von Mädchen übertrafen. Winkelmann et al. (2008) konnten ebenso einen Vorsprung in der allgemeinen Mathematikleistung der Jungen gegenüber den Mädchen in der dritten und vierten Klasse nachweisen. Auch Klieme (1997) bestätigte die frühe Überlegenheit der Jungen in mathematischen Fähigkeiten für den deutschsprachigen Raum.

Wie bereits erläutert, sollten allerdings nicht nur die globalen Durchschnittswerte der erbrachten Leistungen betrachtet, sondern auch die Anteile von Mädchen und Jungen in verschiedenen Leistungsgruppen der Vor- und Grundschule genauer analysiert werden. In einer Studie von Penner und Paret (2008) zeigte sich bei detaillierter Betrachtung der Verteilung der mathematischen Fähigkeiten im letzten Kindergartenjahr, dass Jungen in der leistungsstärksten Gruppe dominierten. Für deutsche Schüler konnte dies aufgrund der IGLU-Studie am Ende der Grundschule bestätigt werden, da hierbei der Anteil der Jungen mit 47% gegenüber einem Anteil von 37% der Mädchen auf den oberen Kompetenzstufen signifikant größer war (Bos et al., 2003).

Auch wenn die Effekte der frühen Fähigkeitsunterschiede teilweise eher gering ausfielen, sind sie doch bedeutend, wenn man die weitere mathematische Entwicklung betrachtet, in deren Verlauf sich die Leistungs- und Fähigkeitsunterschiede zwischen Mädchen und Jungen vergrößern könnten.

1.1.2 Ursachen von Geschlechtsunterschieden in mathematischen Fähigkeiten

Als Erklärung von Geschlechtsunterschieden in mathematischen Fähigkeiten werden neben biologischen und soziokulturellen Ursachen auch andere Fähigkeiten angeführt, die mathematische Leistungen geschlechtsdifferent beeinflussen. So konnten Davis und Carr (2002) unterschiedliche Problemlösestrategien von Mädchen und Jungen nachweisen. Gleichzeitig zeigten verschiedene Untersuchungen, dass Geschlechtsunterschiede unterschiedlich auftraten, je nachdem welche Aufgaben in die Analysen mit einbezogen wurden (Gallagher, Levin, & Cahalan, 2002; Spelke, 2005; Zohar & Gershikov, 2008). Eventuell enthielten mathematische Tests in Studien, in denen Jungen bessere Leistungen zeigten als Mädchen, Aufgaben, für die Jungen bessere Lösungsstrategien nutzen konnten.

Andererseits zeigten sich in mehreren Studien hohe Zusammenhänge zwischen den mathematischen und verbalen Fähigkeiten (Aiken, 1971; Möller, Pohlmann, Köller & Marsh, 2009). Da Mädchen im Allgemeinen bessere verbale Fähigkeiten aufweisen (z.B. OECD, 2010), könnten mögliche Nachteile in den mathematischen Fähigkeiten durch die verbalen Fähigkeiten ausgeglichen werden und tatsächlich vorhandene Geschlechtsunterschiede verborgen bleiben.

Biologische Ursachen der Geschlechtsunterschiede werden von einigen

Autoren dadurch begründet, dass Sexualhormone bereits während der pränatalen Phase unterschiedlich auf die Entwicklung der Gehirnstrukturen von Mädchen und Jungen wirken (Baron-Cohan, 2003; Geary, 1998; Halpern, 2000; Kimura, 2002). Außerdem zeigte Hampson (1990) den Einfluss aktueller Sexualhormonwerte auf visuell-räumliche Fähigkeiten, die mitentscheidend für mathematische Leistungen sind (Schuchardt & Mähler, 2010). Hierbei lösten Frauen visuell-räumliche Aufgaben je nach Phase ihres Menstruationszyklus unterschiedlich gut. Die angesprochene Wirkung von Sexualhormonen erklärt allerdings nicht die mathematischen Geschlechtsunterschiede, die erst ab dem Vorschulalter und noch vor der Pubertät auftreten.

Soziokulturelle Faktoren scheinen als Erklärung für Geschlechtsunterschiede in diesem Altersbereich geeigneter zu sein, denn bereits zweijährige Kinder sind sich ihrer Geschlechtszugehörigkeit bewusst (Huston, 1987) und beginnen auch frühzeitig die Einstellungen und Verhaltensweisen bedeutsamer Personen in ihrer Umwelt diesbezüglich zu übernehmen (Cooley, 1902; Mead, 1934). Da die gesellschaftlichen Annahmen im mathematischen Bereich häufig geschlechtsstereotypisch geprägt sind, also Mädchen im Gegensatz zu Jungen mathematisch als unbegabter angesehen werden, zeigten sich hierbei schon zu Beginn der Grundschulzeit negative Effekte bezüglich der Einstellungen von Mädchen (Ambady, Shih, Kim, & Pittinsky, 2001; Beilock, Gunderson, Ramirez, & Levine, 2010; Lummis & Stevenson, 1990). Dass die negativen Einstellungen zu entsprechenden Leistungsdefiziten führten, ist mit dem mehrfach belegten Zusammenhang zwischen den Einstellungen bezüglich der eigenen akademischen Fähigkeiten (akademisches Selbstkonzept) und den Mathematikleistungen erklärbar (Marsh, 1990; Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006). Das

akademische Selbstkonzept gehört zu den meist untersuchten Erklärungsfaktoren. Deshalb werden im Folgenden die wichtigsten Erkenntnisse bezüglich des akademischen Selbstkonzepts und der entsprechenden Unterschiede zwischen Mädchen und Jungen dargestellt.

1.2 Akademisches Selbstkonzept

Das akademische Selbstkonzept beschreibt die Vorstellungen über die eigenen akademischen Fähigkeiten. Für das Selbstkonzept postulierte James (1890) bereits Ende des 19. Jahrhunderts eine Multidimensionalität, was bedeutet, dass neben einem allgemeinen Selbstkonzept einzelne Selbstkonzeptdimensionen (z.B. soziales Selbstkonzept, körperliches Selbstkonzept) für verschiedene Eigenschaften und Fähigkeiten der Person existieren können. Shavelson, Hubner und Stanton (1976) wiesen die Multidimensionalität des Selbstkonzepts erstmals auch empirisch nach. Auch sie postulierten ein allgemeines Selbstkonzept, welches sich zunächst in ein akademisches (schulisches) und nicht-akademisches Selbstkonzept unterteilt, die sich wiederum aus weiteren Selbstkonzeptdimensionen ergeben (siehe Abbildung 1.1). Marsh, Byrne und Shavelson (1988) widerlegten allerdings die ursprünglich angenommene hierarchische Struktur. Sie konzentrierten sich dabei auf das akademische Selbstkonzept und konnten anhand von Faktor- und Korrelationsanalysen zeigen, dass sich das akademische Selbstkonzept nicht aus allen fachspezifischen Selbstkonzepten im Einzelnen ergibt, sondern aus dem mathematischen und verbalen Selbstkonzept, welche sich wiederum aus unterschiedlichen Anteilen der fachspezifischen Selbstkonzepte zusammensetzen.

besteht, werden die Leistungen nicht direkt für das entsprechende akademische Selbstkonzept übernommen. So werden akademische Selbstkonzepte z.B. aufgrund des Vergleichs mit einer Bezugsgruppe angepasst. Dies wird vor allem beim Big-Fish-Little-Pond-Effekt deutlich (Gabriel, Kastens, Poloczek, Schoreit & Lipowsky, 2010; Marsh, 1987, 2005). Von zwei Personen mit objektiv gleicher Leistung wird diejenige Person ein höheres Selbstkonzept entwickeln, die mehr Möglichkeiten zu Abwärtsvergleichen in der Bezugsgruppe hat, also Teil einer leistungsschwächeren Gruppe ist. Auch bei den temporalen Vergleichen zeigte sich, dass die Person ein höheres akademisches Selbstkonzept entwickelt, die sich in ihren Leistungen steigern konnte (Rheinberg, 2006). Die dimensionalen Vergleiche wurden hauptsächlich in Untersuchungen zum Internal/External-Frame-of-Reference-Modell (Marsh, 1986; Möller, Pohlmann, Köller & Marsh, 2009) angesprochen. Dabei wurde nachgewiesen, dass Schüler ihre fachspezifischen Selbstkonzepte nicht nur durch soziale (externale) Vergleiche ihrer mathematischen und verbalen Leistungen entwickeln, sondern auch durch dimensionale (internale) Vergleiche der Leistungsbereiche untereinander.

Dass die eigenen Leistungen mit dem akademischen Selbstkonzept in Zusammenhang stehen, ist somit gesichert, wobei allerdings zwei verschiedene Wirkrichtungen nachgewiesen wurden (Dickhäuser, 2006; Helmke, 1992). Einerseits wurde u.a. beim bereits erwähnten Big-Fish-Little-Pond-Effekt deutlich, dass die Leistungen die Entwicklung des Selbstkonzepts beeinflussen. Dies wird als Skill-Development-Ansatz bezeichnet (Calsyn & Kenny, 1977; Helmke & van Aken, 1995). Andererseits zeigten weitere Untersuchungen, dass sich auch die Leistungen in Abhängigkeit vom akademischen Selbstkonzept verändern (Trautwein et al., 2006; Valentine et al., 2004). Beim entsprechenden Self-Enhancement-Ansatz

wird angenommen, dass das akademische Selbstkonzept in einem positiven Zusammenhang zum Interesse steht, welches wiederum aufgrund motivationaler Wirkmechanismen die Leistung beeinflusst (Calsyn & Kenny, 1977; Helmke, 1992; Marsh, 1990). Da beide Ansätze mehrfach empirisch bestätigt wurden, bietet das Reciprocal-Effects-Model (Überblick Marsh & Craven, 2006) eine Erklärung für das Zusammenwirken beider Wirkrichtungen. Hierbei wird ein reziproker, dynamischer Zusammenhang zwischen akademischen Leistungen und dem entsprechenden Selbstkonzept angenommen.

1.2.1 Akademisches Selbstkonzept im Vor- und Grundschulalter

Die multidimensionale Struktur des Selbstkonzepts konnte schon bei Kindern ab dem Vorschulalter nachgewiesen werden (Entwisle, Alexander, Pallas & Cadigan, 1987; Harter & Pike, 1984; Helmke, 1991). Wie schon Shavelson et al. (1976) zeigten auch Marsh, Barnes, Cairns und Tidman (1984), dass die einzelnen Selbstkonzeptdimensionen bei Kindern im Vor- und frühen Grundschulalter noch hoch untereinander korrelieren. Allerdings konnten trotzdem einzelne Dimensionen, wie u.a. das verbale, mathematische und akademische Selbstkonzept, separat identifiziert werden. Mit zunehmendem Alter lassen sich diese dann immer besser voneinander trennen (Marsh, Craven & Debus, 1991).

Auffällig bei Kindern im Vor- und Grundschulalter ist, dass sie ihre Fähigkeiten stark überschätzen (Eccles, Wigfield, Harold & Blumenfeld, 1993; Helmke, 1998; Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Kammermeyer & Martschinke, 2006; Nicholls, 1978; Poloczek, Karst, Praetorius & Lipowsky, 2011). Als Grund kann die noch andauernde kognitive Entwicklung angegeben werden, wodurch Kinder erst ab einem Alter von neun Jahren in der Lage sind, relevante

Informationen vollständig zu berücksichtigen und in ihr Selbstkonzept aufzunehmen (Nicholls, 1978; Shavelson et al., 1976; Valentine et al., 2004). Das Selbstkonzept sinkt während der Schulzeit bis ins frühe Erwachsenenalter auf ein realistisches Niveau ab, was in Längsschnittuntersuchungen nachgewiesen wurde (Helmke, 1998; Jacobs et al., 2002).

Es zeigen sich auch schon während der Grundschulzeit Zusammenhänge zwischen Leistungen und dem Selbstkonzept. Hierbei konnten aber kaum Belege für den Einfluss des Selbstkonzepts auf die Leistungen (Self-Enhancement-Ansatz) gefunden werden, was wiederum mit der noch unausgereiften kognitiven Entwicklung begründet wurde (Valentine et al, 2004). Helmke und van Aken (1995) wiesen dagegen den Skill-Development-Ansatz für die Klassenstufen zwei bis vier nach, also den Einfluss von Lehrereinschätzungen in Form von Noten auf die entsprechende akademische Selbstkonzeptdimension. Marsh und Craven (2006) bestätigten dies, wobei sie aber auch betonten, dass sich im Laufe der Grundschulzeit Leistungen und das Selbstkonzept im Sinne des Reciprocal-Effect-Model gegenseitig beeinflussen.

Allerdings werden auch in diesem Altersbereich die Leistungsrückmeldungen nicht direkt für die Selbstkonzeptgenese übernommen. So konnten Gabriel et al. (2010) den Einfluss der mittleren Klassenleistung aufgrund von sozialen Vergleichen bei Kindern der Klassenstufen eins und zwei belegen. Auch Helmke (1991) betonte die Anwendung sozialer Vergleiche bei Grundschulkindern. Dickhäuser und Galfe (2004) fanden weder einen Einfluss dimensionaler noch temporaler Vergleiche auf das akademische Selbstkonzept bei Schülern der dritten und vierten Klasse. Dies konnte auch in der Überblicksarbeit zum Internal/External-Frame-of-Reference-Modell von Möller und Köller (2004)

bestätigt werden, in der sich zeigte, dass dimensionale Vergleiche nicht vor der vierten Klasse durchgeführt werden. Soziale Vergleiche scheinen also hauptsächlich für die Selbstkonzeptgenese von jungen Kindern herangezogen zu werden (Poloczek et al., 2011).

1.2.2 Erfassung des akademischen Selbstkonzepts im Vor- und Grundschulalter

Die üblicherweise angewandten Fragebogenmethoden zur Erfassung des akademischen Selbstkonzepts sind für Kinder im Vor- und frühen Grundschulalter aufgrund mangelnder Lesefähigkeiten nicht geeignet. So wurden in den meisten Untersuchungen mit jungen Kindern interviewgeleitete Befragungen in Einzel- oder kleinen Gruppensettings durchgeführt (Dickhäuser & Stiensmeier-Pelster, 2003; Gabriel et al., 2010; Jacobs et al., 2002; Marsh et al., 1991). Die Antworten der einfach konstruierten Fragen wurden meist zusätzlich mit Unterstützung von Bildern und teilweise durch ein mehrstufiges Antwortverfahren erhoben. Nicholls (1978) und Helmke (1991) wählten in ihren Untersuchungen keine Aussagen oder Fragen, um die Einschätzung der Kinder bezüglich ihrer Fähigkeiten zu erfassen, sondern baten die Kinder, sich in eine Rangfolge von Vergleichskindern einzuordnen, die durch Spielfiguren dargestellt waren. So konnte auf direkte Art und Weise ein Selbstkonzeptwert entnommen werden, wobei kritisch angemerkt werden muss, dass hierbei ein Vergleich mit einer großen und undefinierten Gruppe von Kindern durchgeführt werden musste, was aufgrund des kognitiven Entwicklungsstands der Kinder fehlerhaft ausfallen könnte.

Es gibt auch validierte und normierte Verfahren zur Selbstkonzepterfassung im frühen Kindesalter. Marsh et al. (1984) entwickelten den Self-Description-

Questionnaire (SDQ) zur Erfassung von sieben verschiedenen Selbstkonzeptdimensionen bei Kindern ab fünf Jahren. Auch Harter und Pike (1984) entwickelten die Pictorial Scale zur einfachen und kindgerechten Selbstkonzepterfassung. Allerdings kommen beide Verfahren aus dem englischsprachigen Raum und deutsche Versionen liegen nur für den SDQ-I zur Erfassung des Selbstkonzepts ab acht Jahren vor (Arens, Trautwein & Hasselhorn, 2011) bzw. wurden erst ab der zweiten Klasse eingesetzt (Asendorpf & van Aken, 1993). Zudem können weder in der deutschen Version der Pictorial Scale (Asendorpf & van Aken, 1993) noch im Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen (FEES; Rauer & Schuck, 2004) fachspezifische Selbstkonzepte differenziert erhoben werden, da sie lediglich eine allgemeine Skala zum akademischen Selbstkonzept beinhalten. Validierte Instrumente zur Erfassung des Selbstkonzepts von Kindern vor dem Schuleintritt und in den ersten Schuljahren sind somit für den deutschsprachigen Raum kaum vorhanden.

1.2.3 Geschlechtsunterschiede beim akademischen Selbstkonzept im Vor- und Grundschulalter

In den meisten Untersuchungen zum akademischen Selbstkonzept wurde auch nach Unterschieden zwischen den Fähigkeitseinschätzungen von Jungen und Mädchen geschaut (Asendorpf & van Aken, 1993; Entwistle et al., 1987; Eccles et al., 1993; Gabriel et al., 2010; Helmke & van Aken, 1995; Jacobs et al., 2002; Marsh et al., 1984; 1991; Valeski & Stipek, 2001; Wigfield & Karpachian, 1991). Im Vor- und Grundschulalter konzentrierten sich die Studien hauptsächlich auf Selbstkonzeptunterschiede im mathematischen und verbalen Bereich. Für das

mathematische Selbstkonzept konnten Hyde, Fennema, Ryan, Frost und Hopp (1990) in einer Metaanalyse zeigen, dass Jungen zwischen fünf und zehn Jahren ihre Fähigkeiten bereits signifikant besser einschätzten als Mädchen. Allerdings ist zu erwähnen, dass der Geschlechtseffekt mit $d = .08$ als sehr gering zu betrachten ist und erst mit zunehmenden Alter ansteigt. Auch Marsh (1989) berichtet eine eher geringe Varianzaufklärung der Selbstkonzeptunterschiede durch das Geschlecht. So wäre auch zu erklären, dass in einigen Studien für diesen Altersbereich keinerlei Unterschiede in den mathematischen Fähigkeitseinschätzungen gefunden wurden (Herbert & Stipek, 2005; Marsh et al., 1991; Valeski & Stipek, 2001). Die meisten Untersuchungen zu Geschlechtsunterschieden im mathematischen Selbstkonzept kommen allerdings zu dem Ergebnis, dass Jungen im Vor- und frühen Grundschulalter ihre Fähigkeiten höher einschätzten als Mädchen (Gabriel et al., 2010; Jacobs et al., 2002; Marsh et al., 1998; Tiedemann & Faber, 1995; Wagner & Valtin, 2003; Wigfield et al., 1997).

Für den Bereich des verbalen Selbstkonzepts ist die Ergebnislage etwas kontroverser. Zum einen gibt es Studien, die zeigen, dass Mädchen ab dem Vorschulalter konstant bis in die frühen Grundschuljahre ein höheres verbales Selbstkonzept aufwiesen als Jungen (Marsh et al., 1984; 1991; Valtin, Wagner & Schwippert, 2005; Wigfield et al., 1997). Andere Studien konnten dagegen keine Selbstkonzeptunterschiede zwischen Mädchen und Jungen berichten (Eccles et al., 1993; Nicholls, 1978; Valeski & Stipek, 2001). In der Längsschnittuntersuchung von Jacobs et al. (2002) zeigte sich, dass zu Beginn der Grundschule Mädchen und Jungen die gleichen Einschätzungen bezüglich ihrer verbalen Fähigkeiten abgaben. Im Laufe der Schuljahre entwickelte sich ein Vorteil für die Mädchen, der allerdings erst ab der Sekundarstufe signifikant wurde. Entgegen den

geschlechtsstereotypischen Erwartungen konnten Wagner und Valtin (2003) sogar einen Vorteil für Jungen der zweiten Klasse im verbalen Selbstkonzept nachweisen, der sich allerdings in den folgenden Schuljahren nicht mehr zeigte.

Beim allgemeinen akademischen Selbstkonzept zeigten sich in den meisten Studien keine Unterschiede zwischen Mädchen und Jungen im Vor- und frühen Grundschulalter (Jacobs et al., 2002; Marsh et al., 1984; 1991). Dies kann mit den oben genannten Ergebnissen begründet werden, da die Vorteile der Jungen im mathematischen Bereich oftmals durch die höheren Selbstkonzepte der Mädchen im verbalen Bereich ausgeglichen wurden.

Die inkonsistenten Ergebnisse bezüglich der Geschlechtsunterschiede im akademischen Selbstkonzept können auf zwei Faktoren zurückgeführt werden. Zum einen scheint z.B. das Konstrukt des verbalen Selbstkonzepts unterschiedlich definiert und erhoben zu werden. Meistens werden die Kinder nach ihren Fähigkeiten im Lesen befragt, es finden sich aber auch Studien, wie die von Wagner und Valtin (2003), in denen das Selbstkonzept für die Rechtschreibfähigkeiten erhoben wird. Beim SDQ-I wird auch nach dem Leseselbstkonzept gefragt, beim SDQ-II allerdings wird das verbale Selbstkonzept über die Einstellung zum muttersprachlichen Unterricht erhoben (Marsh, 1989; Marsh et al., 1991). Ähnlich verhält es sich im mathematischen Bereich. Fünf- bis achtjährige Kinder sollen beim SDQ-IA ihre Einschätzungen zum Umgang mit Zahlen angeben; beim SDQ-I hingegen werden die Einschätzungen der älteren Grundschulkinder zum Mathematikunterricht erhoben. Die zweite mögliche Ursache der unterschiedlichen Ergebnisse könnte in der variierenden Anzahl und der Art an Items zur Erfassung des akademischen Selbstkonzepts begründet sein. In den bisher berichteten Untersuchungen wurden zwischen ein bis zwei (Nicholls, 1978; Valeski & Stipek,

2001; Wagner & Valtin, 2003) und zehn Items (Eccles et al., 1993; Marsh 1989) zur Selbstkonzepterfassung im jeweiligen Bereich genutzt. Wurden dann entweder nur affektive (z.B. Wie sehr magst du Lesen?; Eccles et al., 1993) oder nur kognitive (z.B. Wie gut bist du im Rechnen?; Wagner & Valtin, 2003) Einschätzungen oder eine Mischung aus beiden genutzt, ist die Vergleichbarkeit der einzelnen Untersuchungsergebnisse stark eingeschränkt.

1.3 Forschungsfragen

Basierend auf den theoretischen Ausführungen ergeben sich mehrere Forschungsdesiderata. So zeigte sich, dass vor allem im deutschsprachigen Raum kaum Untersuchungen zu mathematischen Leistungen und entsprechenden Geschlechtsunterschieden vor Mitte der Grundschulzeit vorhanden sind. Die Studien, die mathematische Fähigkeiten im Vor- und frühen Grundschulalter erhoben haben, liefern inkonsistente Ergebnisse bezüglich der Leistungsunterschiede zwischen Mädchen und Jungen. Zusätzlich wurden kaum mögliche Einflussfaktoren bei der Betrachtung von mathematischen Leistungsunterschieden zwischen Mädchen und Jungen einbezogen. Hierfür bieten sich einerseits die verbalen Fähigkeiten und andererseits das akademische Selbstkonzept an. Auch beim akademischen Selbstkonzept fällt auf, dass die Erhebung im Vor- und frühen Grundschulalter eher selten durchgeführt wird. Zusätzlich scheinen die bisher verwendeten Erhebungsmethoden nicht ganz passend für diese junge Altersgruppe zu sein, da die Selbstkonzeptwerte von Vor- und Grundschülern positiv verzerrt sind. Schilling, Sparfeldt und Rost (2006) betonen:

„Bislang haben sich leider nur wenige Arbeiten mit einer umfassenderen Analyse fachspezifischer Ske [Selbstkonzepte] befasst (...), allerdings liegen kaum Befunde zum Einfluss des Geschlechts vor.“ (Schilling et al., 2006, S. 16)

Hier besteht also weiterer Forschungsbedarf, vor allem im Vor- und Grundschulbereich. Somit ergeben sich für die hier vorliegende Arbeit folgende Fragestellungen:

1. Gibt es mathematische Leistungsunterschiede zwischen Mädchen und Jungen im Vor- und frühen Grundschulalter?
2. Treten Veränderungen in mathematischen Leistungsunterschieden zwischen Mädchen und Jungen vom Vor- bis zum frühen Grundschulalter auf?
3. Haben verbale Fähigkeiten Einfluss auf mathematische Leistungen und Geschlechtsunterschiede in Mathematik?
4. Lassen sich einzelne Bereiche des akademischen Selbstkonzepts bei Kindern im frühen Grundschulalter mit einer neuen Erhebungsmethode valider erfassen als mit den üblicherweise verwendeten Fragebogenmethoden?
5. Zeigen sich Unterschiede im akademischen Selbstkonzept zwischen Mädchen und Jungen im frühen Grundschulalter bei Verwendung unterschiedlicher Methoden zur Selbstkonzepterfassung?

In den folgenden Kapiteln dieser Arbeit werden drei Studien vorgestellt, die sich mit den genannten Fragestellungen befassen. Die erste Studie beschäftigt sich mit Geschlechtsunterschieden in mathematischen und verbalen Fähigkeiten von Kindern im Vor- und frühen Grundschulalter und gibt somit Antworten zu den

Fragestellungen eins bis drei. In der zweiten Studie wird der Mitschülervergleich als neue Erhebungsmethode für das akademische Selbstkonzept zu Beginn der Grundschule vorgestellt, wodurch die vierte Forschungsfrage beantwortet werden kann. In der dritten Studie werden Geschlechtsunterschiede beim akademischen Selbstkonzept von Schülern im frühen Grundschulalter untersucht, so dass Antworten zur fünften Fragestellung gefunden werden können. Im Anschluss an die Darstellung der Studien werden die Ergebnisse im abschließenden Kapitel zusammengefasst und diskutiert.

2. Studie I²

Girls and math – a never-ending story!?

Abstract

The discussion whether there are gender differences in math abilities is both heated and seemingly never-ending. Whereas some studies gave evidence for a lack of gender differences, a lot of research revealed such differences even in elementary school children. Because most of the studies did not investigate children before the middle of elementary school, this paper tries to close the gap in focusing on how girls and boys from kindergarten to 2nd grade differ in their math abilities. Our results support findings from studies that showed gender differences existing already at that very young age. Boys outperformed girls in math when taking verbal abilities into account. These results can be used to take a deeper look on gender issues concerning math education especially in Germany.

Key words: gender differences, math ability, kindergarten children, elementary school children

2.1 Introduction

Gender differences in math and science are heatedly discussed (e.g., Else-Quest, Hyde, & Linn, 2010; Halpern et al., 2007; Hosenfeld, Köller, & Baumert, 1999; Hyde, Fennema, & Lamon, 1990; Lachance & Mazzocco, 2006; Mullis et al., 1997;

² Based on a manuscript submitted for publication: Sahr, K., Arndt, D., Opfermann, A., Fritz, A., Krinzinger, H., Willmes, K., & Leutner, D. (2012) *Girls and math – A never-ending story!?*. Manuscript submitted for publication.

OECD, 2004, 2006, 2009; Penner, 2008; Spelke, 2005; Winkelmann, van den Heuvel-Panhuizen, & Robitzsch, 2008). The discussion of differences in math abilities between girls and boys does not take place just for its own sake, but is highly important in the current discussion about the societal need for qualified employees in science, technology, engineering, and math (commonly abbreviated as STEM). In 2008 the Cologne Institute for Economic Research (IW), commissioned by the German Federal Ministry of Economics and Technology, pointed out that there is a lack of qualified specialists in the field of math and science. With the ongoing development this gap might even increase in the next years. It is recommended by the authors, that increasing the number of graduates in STEM fields of study from currently 32% to at least 40% of all university graduates could be the answer to this problem. Because the proportion of female graduates in these latter fields of study is only 25% (OECD, 2006, 2008) and the percentage of women, who do their apprenticeship in the field of computer engineering, is very low as well (Struwe, 2005), there is a high female potential of future STEM specialists.

The low rates of young women in math or technical jobs seem to have their origin in typical school careers of girls and women (Halpern et al., 2007; Hannover, 1991). Heine, Kerst, and Sommer (2007) gave an explanation for this finding by revealing high correlations between the choice of advanced courses in school and the following field of study. In addition, it was shown that female students in secondary school STEM classes prefer basic rather than advanced courses in contrast to male students, who choose advanced courses by majority (Hannover, 1991; Köller, Daniels, Schnabel, & Baumert, 2000; Sherman & Fennema, 1977). Thus, the cause for later career choices has to be located to a large degree in times of school or even earlier. Thus, it is necessary to pay attention to gender differences in

math abilities over the whole educational period to identify the point of time at which differences start to occur. However, so far only few studies have been focusing on math abilities of children in kindergarten or first elementary school years - especially in the German-speaking countries. Thus, the aim of our study was to find out whether there are differences concerning math abilities in children from kindergarten to 2nd grade and, if so, how these differences develop over this period of time.

2.2 Theoretical background

Gender differences in math

Research findings concerning gender differences in math abilities are still inconsistent. Mainly North-American studies indicated that differences between girls and boys have become smaller over time. For instance, in their meta-analysis, Hyde et al. (1990) showed that gender differences had decreased from $d = .31$ (Cohen's d , 1969), based on studies between 1967 and 1973, to an average effect size of $d = .14$ in studies from 1974 to 1987. A more recent analysis of the annual achievement tests of US students from 2nd to 11th grade revealed average effect sizes ranging from $d = -0.02$ to $d = 0.06$ (in favor of male students) which are quite small (Hyde et al., 2008). In another meta-analysis of international PISA and TIMSS 2003 data, Else-Quest et al. (2010) supported this finding in the way that the average effect sizes they found for gender differences regarding math were small as well ($d < 0.15$).

Analyses of PISA and TIMSS data showed marginal gender differences across all participating countries (OECD, 2004, 2009, 2010). But concerning the size of gender differences within countries, large differences between countries

could be found. For instance, Zimmer et al. (2004) found that in South Korea boys outperformed girls in PISA 2003 by reaching 23 points more in the math test ($d = 0.26$), whereas, in contrast, Icelandic girls performed significantly better than their male counterparts ($d = -0.17$). On the other hand, looking at the math competence levels of the PISA 2003 study, it can be seen that, across all countries, 17% of male students and only 12% of female students reached the two highest levels, whereas the allocation of boys and girls at the two lowest levels (21% of boys vs. 22% of girls) was counterbalanced (OECD, 2004). Special analyses of TIMSS data confirmed this finding in so far as in all participating countries boys outperformed girls at the top of the distribution (90th percentile) of math abilities (Penner, 2008). As can be seen in Table 2.1, the superiority of German boys over German girls remained significant in all PISA studies from 15 points in the year 2000 to 16 points in 2009 (OECD, 2004, 2009, 2010; Stanat & Kunter, 2001). Thus, we cannot talk of decreasing gender differences in math in Germany.

Furthermore, in contrast to Hyde et al. (1990), a meta-analysis of studies focusing on gender differences in math revealed that the superiority of 16-to 30-year-old men was larger in German-speaking countries than in North America (d from 0.49 to 0.55; Klieme, 1997). In addition, no decline of gender differences could be found over the years. Thus, it is interesting to know if and to what extend girls and boys differ in their math abilities in earlier years of age as well.

Early math abilities and respective gender differences

A remarkable amount of studies was conducted to explore basic numerical and math abilities of babies right after birth and in their first months of life (e.g., Baillargeon, 1993; Baillargeon & DeVos, 1991; Barth et al., 2003;

Table 2.1: Mean score and gender differences in student performance on the PISA math scale (adapted from OECD, 2004, 2009, 2010; Stanat & Kunter, 2001)

Countries	2000			2003			2006			2009		
	boys	girls	diff	boys	girls	diff	boys	girls	diff	boys	girls	diff
Australia	539	527	12	527	522	5	527	513	14	519	509	10
Austria	530	503	27	509	502	8	517	494	23	506	486	19
Belgium	524	518	6	533	525	8	524	517	7	526	504	22
Canada	539	529	10	541	530	11	534	520	14	533	521	12
Chile							424	396	28	431	410	21
Czech Republic	504	492	12	524	509	15	514	504	11	495	490	5
Denmark	522	507	15	523	506	17	518	508	10	511	495	16
Estonia							515	514	1	516	508	9
England	534	526	9									
Finland	537	536	1	548	541	7	554	543	12	542	539	3
France	525	511	14	515	507	9	499	492	6	505	489	16
Germany	498	483	15	508	499	9	513	494	20	520	505	16
Greece	451	444	7	455	436	19	462	457	5	473	459	14
Hungary	492	485	7	494	486	8	496	486	10	496	484	12
Iceland	513	518	-5	508	523	-15	503	508	-4	508	505	3
Ireland	510	497	13	510	495	15	507	496	11	491	483	8
Israel							448	436	12	451	443	8
Italy	462	454	8	475	457	18	470	453	17	490	475	15
Japan	561	553	8	539	530	8	533	513	20	534	524	9
Korea	559	532	27	552	528	23	552	543	9	548	544	3
Luxembourg	454	439	15	502	485	17	498	482	17	499	479	19
Mexico	393	382	11	391	380	11	410	401	9	425	412	14
Netherlands				540	535	5	537	524	13	534	517	17
New Zealand	536	539	-3	531	516	14	527	517	11	523	515	8
Norway	506	495	11	498	492	6	493	487	6	500	495	5
Poland	472	468	5	493	487	6	500	491	9	497	493	3
Portugal	464	446	19	472	460	12	474	459	15	493	481	12
Slovak Republic				507	489	19	499	485	14	498	495	3
Slovenia							507	502	5	502	501	1
Spain	487	469	18	490	481	9	484	476	9	493	474	19
Sweden	514	507	7	512	506	7	505	500	5	493	495	-2
Switzerland	537	523	14	535	518	17	536	523	13	544	524	20
Turkey				430	415	15	427	421	6	451	440	11
United Kingdom	534	526	8				504	487	17	503	482	20
United States	497	490	7	486	480	6	479	470	9	497	477	20
OECD total				494	484	10	489	478	12	496	481	15
OECD average	506	495	11	506	494	11	503	492	11	501	490	12

Note: Values that are statistically significant are indicated in bold.

Feigenson & Carey, 2003; Feigenson et al., 2002; Hespos & Baillargeon, 2001;

Kellman, 1993; Spelke et al., 1995; van Oeffelen & Vos, 1982; Wynn, 1992, 1995;

Xu & Arriaga, 2007; Xu & Spelke, 2000). These studies revealed that even very

young infants seem to possess innate principles which allow them to discriminate

between two sets of objects and to solve simple addition and subtraction problems.

In addition, these early numerical abilities provide the basis for the later development of math concepts.

Studies concerning math development analyzed the acquisition of central arithmetical concepts in the following years of children's life, starting with the ordinal and cardinal understanding of numbers (Brannon & Van de Walle, 2001; Le Corre, Van de Walle, Brannon, & Carey, 2006; Wynn, 1990), the ability to operate with bigger amounts (Barth et al., 2006), until the understanding of the part-part-whole principle at the age of about seven (Piaget, 1965; Steffe, 1992).

Spelke (2005) emphasized in a paper on gender differences in math that no substantial differences between young girls and boys (ranging from 6 months to 7 years) were found in any of her observations. In kindergarten and elementary school some studies demonstrated that boys and girls achieved the same results even in more complex tasks (Hyde et al., 1990; Lachance & Mazzocco, 2006). In a longitudinal study, Tiedemann and Faber (1994) compared math test scores and teacher ratings from preschool until the end of elementary school in grade four. They could not find any gender differences as well. In contrast to these findings, other studies showed that young boys outperformed their female counterparts. Weinhold Zulauf et al. (2003) found that seven-year-old boys in Swiss kindergartens outperformed girls of the same age. Six months before, girls in that study had performed better than boys in basic math abilities, but this difference did not reach statistical significance. In the Netherlands, elementary school boys outperformed girls constantly in the five-year MOOJ study (van den Heuvel-Panhuizen, 2004). Studies from Germany supported the boys' early predominance in math abilities as well (Klieme, 1997). In TIMSS 2007, there were no gender differences between 4th grade boys and girls in overall mean scores across countries, but in Germany boys

outperformed girls significantly ($d = 0.12$; Bos et al., 2008). Winkelmann et al. (2008) examined gender differences for grade three and grade four children. The average mean scores of boys were higher than those of the girls, although boys did not dominate in all kinds of math problems. In the Annual German Comparative Tests in elementary schools (VERA 4; Helmke et al., 2005; 2006), boys achieved significantly better results for story problems with effect sizes of $d = 0.30$ (2004) and $d = 0.29$ (2005).

Not only overall mean scores are important to be analyzed, but the distribution of boys' and girls' math achievement has to be taken into consideration as well. Penner and Paret (2008) found an advantage for boys at the top of the distribution in the last year of kindergarten. But not until the end of first grade boys achieved significantly better results than girls on average. Analyses of the German national extension of the Progress in International Reading Literacy Study (PIRLS; Walther, Geiser, Langeheine, & Lobemeier, 2003) showed that more boys than girls reached the two highest math competency levels (46% versus 37%) and that there were more girls than boys in the two lowest competency levels (22% versus 15%). At the medium competency level, the number of boys and girls was counterbalanced (39% versus 41%).

Causes of gender differences

Possible reasons for math ability differences between girls and boys are manifold and controversially discussed. Some researchers claimed a biological influence to be responsible for gender differences (Baron-Cohan, 2003; Geary, 1998). Their argument is that sex hormones have a proven influence on the development of neuronal structures in prenatal development so that there are differences in girls' and

boys' neuroanatomy of the brain (Halpern, 2000; Kimura, 2002). But if focusing on this fact only, the question would arise why no differences in math abilities were found with children from birth until six years of age. In contrast to that, the influence of sex hormones on cognitive abilities was demonstrated by Hampson (1990) who proved that women achieved differently in visual-spatial tasks regulated by their menstruation cycle. However this reason does not count for different math abilities in children before adolescence.

In fact, not only genetic or hormonal factors have been considered to be responsible for human behavior but also environmental circumstances, which are expected to have an important impact on the development of neuronal structures and in turn on behavior (Roffman, Marci, Glick, Dougherty, & Rauch, 2005). Children of two years of age are already aware of their gender classification (Huston, 1987) and they soon start to take over attitudes and behavior patterns from relevant persons. In the field of math, this copying effect often has negative implications on achievement results and, respectively, on the self-concept of girls (Ambady et al., 2001; Beilock et al., 2010; Lummis & Stevenson, 1990).

Furthermore, different usage of problem solving strategies was shown for male and female children (Davis & Carr, 2002). This finding presents a possible explanation for the controversial results of studies focusing on gender differences in math abilities. Perhaps boys performed better in some studies because math problems required strategies boys liked more (Gallagher et al., 2002; Zohar & Gershikov, 2008). In studies using the Scholastic Assessment Test (SAT-M), gender differences emerged (boys outperformed girls) after some tasks had been excluded from the data analyses (Spelke, 2005). Thus, it seems that gender differences appear differently with different types of math problems.

In addition, verbal abilities or phonological awareness are important for arithmetic calculation. Berg (2008) revealed an influence of verbal abilities on math achievement. Krajewski and Schneider (2009) supported these findings with correlations of verbal and math abilities, ranging from $r = .47$ to $r = .51$. Furthermore, they showed an influence of early phonological awareness in kindergarten on basic numerical abilities in 1st grade as well. In a meta-analysis, Möller, Pohlmann, Köller, and Marsh (2009) found an average correlation between verbal and math abilities of $r = .67$ out of 69 studies with more than 125,000 participants. They also showed that the correlation is larger when looking at test results rather than grades ($r = .74$ vs. $r = .54$). All these correlations cannot be explained by general intelligence only. In a review of studies from the 1950s and 1960s, Aiken (1971) revealed that the correlation between verbal and math abilities, ranging from $r = .40$ to $r = .86$, became smaller when intelligence was partialled out. But even the smaller correlations remained statistically significant ($r = .13$ to $r = .46$). Girls have an advantage in the domain of verbal abilities (e.g. OECD, 2010). Gender differences in math could be covered because girls compensate their low math abilities with higher verbal abilities. Thus, gender differences in math results could become apparent as recently as controlling for verbal abilities.

Research questions

Apart from the fact that math differences in favor of boys in their early school career were often found to be rather small, they are nevertheless important regarding further math development, because even small differences in knowledge and skills can increase over time (Schwippert, Bos, & Lankes, 2003). Based on the previously described studies, it is evident that only few studies exist which focus on gender

differences in math abilities of children in kindergartens and elementary schools. In particular, research studies focusing on math abilities typically begin in the middle of elementary school. But at this point of time, differences between girls and boys might already be manifest (Bos et al., 2003; Klieme, 1997; Winkelmann et al., 2008). Only few studies dealt with younger children, and the respective results seem to be inconsistent (Tiedemann & Faber, 1994; Weinhold Zulauf et al., 2003). Thus, there is still a lack of data on possible gender differences.

In addition, in the TIMS study in 1995, three different age populations were tested at the same time (Mullis et al., 1997). The effect sizes of gender differences in math abilities increased with increasing age. In the population of the nine-year-olds the effect was marginal ($d = 0.02$) but it became larger in the population of the 13-year-olds ($d = 0.08$) and increased to $d = 0.33$ for the 16-year-olds (see Burba, 2006; Winkelmann et al., 2008). Penner and Paret (2008) found different gender effects within different age populations in preschool and elementary school children as well. If gender differences can be assumed in general, the question still remains, how much young girls and boys differ in their math abilities. Thus, our first research question is as follows:

- 1) Does a gender difference in math abilities already exist in kindergarten and does this difference increase until 2nd grade?

Possible reasons for girls' and boys' different test results are discussed above. For example, it was mentioned that verbal abilities play an important role in math achievement. Thus, finding no gender differences might possibly be traced back to math problems for which high verbal abilities are needed. Then boys, who perform worse regarding verbal abilities, might not be able to show their math expertise.

Thus, verbal abilities have to be taken into account when analyzing math test results of girls and boys. Accordingly, our second research question was as follows:

- 2) Are differences in verbal abilities related to differences in math achievement results?

2.3 Method

Sample & Procedure

In the present study, 88 children were tested for their math, verbal, and cognitive abilities. Their mean age was $M = 77.4$ months ($SD = 16.5$). The number of girls and boys was almost counterbalanced (42 vs. 46). The boys and girls attended regular kindergartens and elementary schools (1st & 2nd grade) in five large German cities.

The testing of the children took place in individual sessions. Math ability was assessed with a test based on a stage model of early arithmetic development (Fritz & Ricken, 2008). All children worked on the same 45 math problems. Elementary school children received 16 additional items, which were not administered to the preschoolers. All 61 items were Rasch scalable, making it possible to measure all children's math abilities on the same scale. The WLE person-separation reliability of this instrument was .945. Verbal abilities were assessed with the scale "Imitation of Grammatical Structures" of the German "Heidelberger Sprachentwicklungstest" (HSET; Grimm & Schöler, 1978) and the phonological awareness scales "Syllabifying" and "Rhyming" of the German "Bielefelder Screening" instrument (BISC; Jansen, Mannhaupt, Marx, & Skowronek, 2002). Additionally, cognitive ability was measured with the matrices subscale of the "Culture Fair Intelligence Test" (CFT1; Cattell, Weiß, & Osterland, 1997). For further analysis factor scores were calculated for all scales of verbal ability together and for the cognitive ability

subscale. All children met age-appropriate standards, that is, they neither had math nor cognitive deficits ($20 \leq \text{age-adjusted percentile rank} \leq 80$).

2.4 Results

To examine whether gender differences in math abilities exist, a two-factorial analysis of variance (ANOVA) was calculated with math achievement as dependent and gender and grade (kindergarten, 1st grade, 2nd grade) as independent variables. The results revealed a main effect for grade, $F(2,82) = 211.24$, $p < .001$, *part. η^2* = .84, and for gender, $F(1,82) = 4.04$, $p = .048$, *part. η^2* = .05. The statistical interaction of grade and gender was not significant, $F(2,82) = 1.11$, $p = .334$, *part. η^2* = .03. Math achievement increases strongly from kindergarten to 2nd grade, and boys always scored higher than girls (Figure 2.1). Even in 2nd grade neither a boy (maximum score of 59) nor a girl (maximum score of 60) solved all 61 tasks correct. Mean scores and standard deviations for all measures are given in Table 2.2.

Table 2.2: Number of boys and girls, math ability (raw scores), verbal ability and residuals (math ability controlled for verbal ability): Means and standard deviations

			Math ability <i>M (SD)</i>	Verbal ability <i>M (SD)</i>	Residuals <i>M (SD)</i>	Cognitive ability <i>M (SD)</i>
<i>n</i>						
Kindergarten	boys	18	21.9 (8.1)	-0.5 (1.1)	-10.0 (7.9)	-0.6 (0.6)
	girls	14	20.6 (9.2)	-0.9 (1.2)	-6.7 (10.7)	-0.6 (0.6)
1st Grade	boys	12	50.5 (6.9)	0.9 (0.4)	3.7 (7.9)	0.7 (0.7)
	girls	11	44.3 (3.4)	0.5 (0.9)	1.9 (8.9)	0.5 (0.8)
2nd Grade	boys	16	53.6 (3.7)	0.5 (0.5)	10.7 (5.8)	1.1 (0.5)
	girls	17	51.8 (4.9)	1.1 (0.4)	2.2 (4.8)	1.2 (0.5)

Contrast analyses show that in 1st grade boys outperformed girls significantly, $F(1,82) = 5.31, p = .024, part. \eta^2 = .06$. This difference results in a large effect of $d = 0.96$. But even the non-significant differences between girls and boys in kindergarten and 2nd grade represent small effect sizes ($d_{kindergarten} = .20, d_{second\ grade} = .28$).

The second research question concerned the relationship of math and verbal abilities. The correlation of both variables, while controlling for cognitive ability, was $r = .60, p < .001$. This result shows that math ability is closely related to verbal ability even when taking general cognitive ability into account, which can explain a large amount of variance of both variables ($r_{math, cognitive} = .85; r_{verbal, cognitive} = .58, p < .001$).

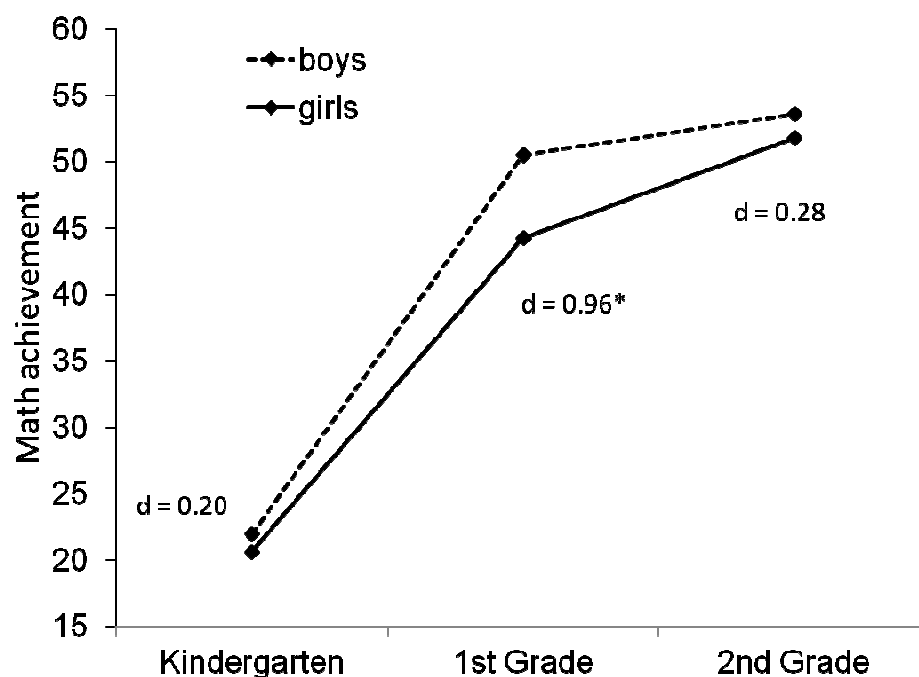


Figure 2.1: Math achievement (total scores) of boys and girls from kindergarten to 2nd grade; * $p < .05$

Looking at the development of verbal abilities from kindergarten to 2nd grade shows a main effect of grade, $F(2,82) = 33.41, p < .001, \text{part. } \eta^2 = .45$, but no main effect of gender, $F(1,82) < 1$. The interaction of gender and grade, however, is significant, $F(2,82) = 4.27, p = .017, \text{part. } \eta^2 = .09$. As can be seen in Figure 2.2 (see also Table 2.2) girls' and boys' verbal development is the same from kindergarten to 1st grade. But while girls' verbal abilities increase continuously in 2nd grade ($t(12.24) = -2.285, p = .041$, t-test for different variances), boys' verbal abilities even decrease ($t(26) = 2.087, p = .047$). The difference between girls' and boys' verbal abilities, computed as a linear contrast, is significant in 2nd grade, $F(1,82) = 4.91, p = .029, \text{part. } \eta^2 = .06, d = .77$.

Taking into account the correlation of verbal and math abilities as well as gender differences in the development of both, a two-factorial ANOVA was

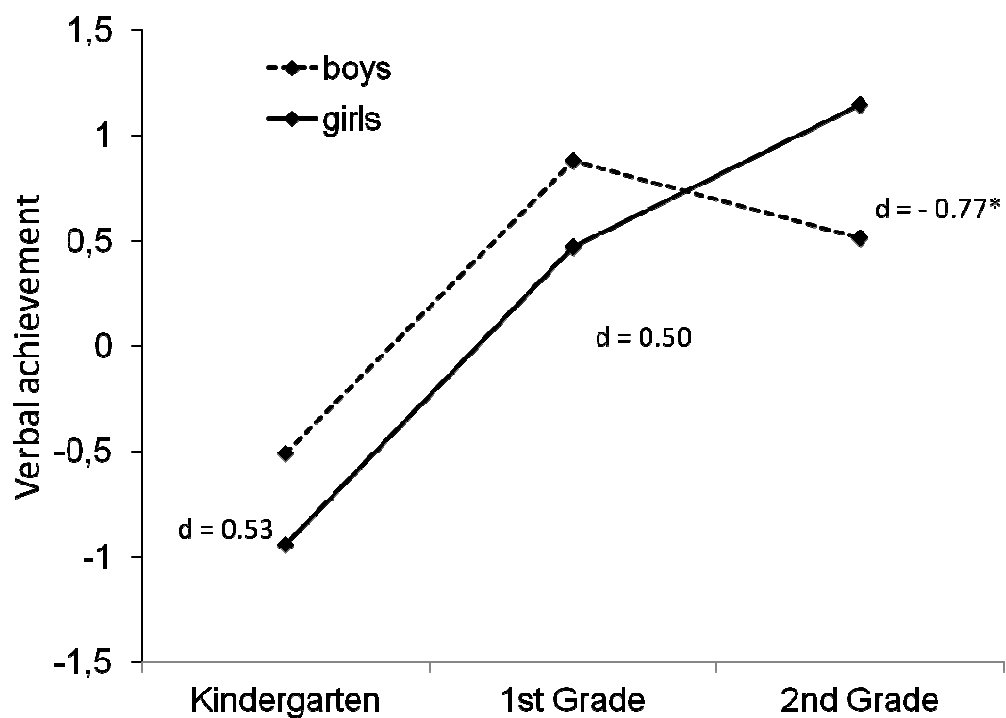


Figure 2.2: Verbal achievement of boys and girls from kindergarten to 2nd grade; * $p < .05$

calculated on the residuals of math achievement after partialling out verbal abilities via simple regression, $b = .75$, $t(86) = 10.46$, $p < .001$; $R^2 = .56$, $F(1,86) = 109.49$, $p < .001$. This analysis revealed a main effect of grade, $F(2,82) = 32.35$, $p < .001$, $part. \eta^2 = .44$, but not of gender, $F(1,82) = 2.31$, $p = .13$, $part. \eta^2 = .03$. Furthermore, a significant interaction of grade and gender indicated differences in the math development of boys and girls from kindergarten to 2nd grade, $F(2,82) = 4.81$, $p = .011$, $part. \eta^2 = .11$. Thus, when controlling for verbal abilities, girls performed better than boys in kindergarten (Figure 2.3 and Table 2.2). This advantage of girls changed with boys achieving better results than girls in 1st grade and 2nd grade. Linear contrast analyses revealed that the differences between boys and girls reached statistical significance only in 2nd grade, $F(1,82) = 10.12$, $p = .002$, $part. \eta^2 = .11$, $d = 1.11$. But even the non-significant differences between girls and boys in kindergarten ($d = -0.43$) and 1st grade ($d = 0.24$) resulted in small to medium effect sizes.

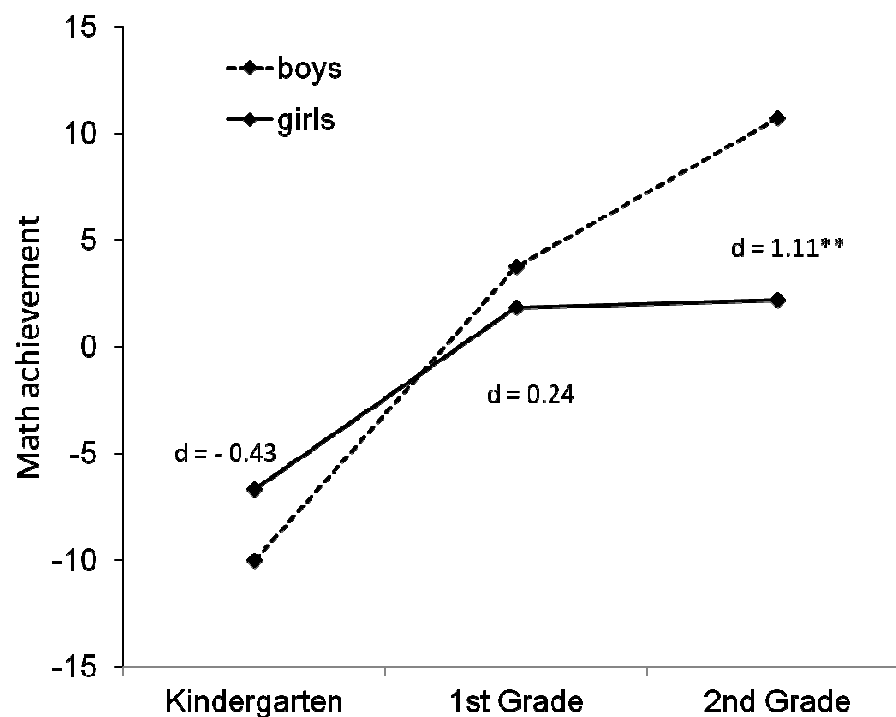


Figure 2.3: Residuals of math achievement (controlled for verbal ability) for boys and girls from kindergarten to 2nd grade; ** $p < .01$

2.5 Discussion

The aim of our study was to find out whether gender differences in math abilities exist for children from kindergarten to 2nd grade. Some studies and respective meta-analyses revealed gender differences in math not until secondary school (Else-Quest et al., 2010; Hyde et al., 1990), these differences being small and remaining small even for older students. On the other hand, some researchers found math differences for girls and boys of younger ages (van den Heuvel-Panhuizen, 2004; Winkelmann et al., 2008). But most of these studies started observing children from 3rd grade on. Gender differences were found in the way that boys outperformed girls regarding the overall math ability. But there is less knowledge about how boys and girls perform before and at the beginning of elementary school. Analyses of the present study revealed that boys significantly outperform their female counterparts in 1st grade but not in kindergarten or 2nd grade. A ceiling effect can be excluded because even in 2nd grade no child reached the maximum score of 61. Thus, the development of gender differences in math abilities seems to be inconsistent. But even though statistical significance of differences was not reached in kindergarten and 2nd grade, the effect sizes of gender differences in math ability in the present study were even larger than the significant effect size of $d = 0.12$ for 4th graders in TIMSS 2007.

This indicates some practical relevance of gender differences in math at all grades, but even more important, from the beginning of formal education.

Furthermore, our results are consistent with the findings of Penner and Paret (2008) in the way that they revealed an overall advantage already for boys in 1st grade. On the other hand, we only studied age-appropriately developed children. Penner and Paret (2008) did not find gender differences with age-appropriately developed children in these grades. But the difference appeared at the end of 1st grade,

increased in 3rd and 5th grade and there, the largest differences could be found in the age-appropriate achievement group. These findings are in line with our results after controlling for verbal abilities. We found that girls and boys achieved the same scores in kindergarten and 1st grade. But in 2nd grade, differences in math abilities increased and boys outperformed girls significantly. Winkelmann et al. (2008) as well as the German PIRL study (Walther et al., 2003) revealed an advantage for boys in 3rd and 4th grade, a finding that fits well with our results. In other words, the findings of Winkelmann et al. (2008) and Walther et al. (2003) also contribute to the discussion of increasing gender differences through the course of young childrens' development. Furthermore, our results support the findings of Weinhold Zulauf et al. (2003) that kindergarten girls had a small but non-significant advantage in math, which shifted to an advantage for boys shortly afterwards. Differences in time of appearance of the boys' predominance are possibly caused by the fact that children with accelerated math abilities as well as children with deficits in math were not included in our study. In addition, different instruments were used to assess math abilities. Thus, verbal abilities may be more important for boys in our study than in other studies. Nevertheless, although statistical significance could not always be reached in our study, small effect sizes still indicate a gender gap in math abilities from the first year of elementary school on.

In summary, on the basis of our results, we cannot approve the assumption of disappearing gender differences in math, as Hyde et al. (1990) had mentioned. In the future, variables such as verbal ability and achievement group might be worth taken into account as potentially influential variables with regard to the comparison of math achievement results.

Some limitations have to be mentioned for our study as well. First, because

of the fact that we only worked with children from Germany, no general implication for math development of children worldwide can be made. The large differences between participating countries in PISA (e.g. OECD, 2009) concerning the gender gap in math abilities are in line with the gender stratification hypothesis (Guiso, Monte, Sapienza, & Zingales, 2008; Else-Quest et al., 2010). This hypothesis ascribes the cross-national patterns in mathematical gender differences to the gender status within the respective culture. This leads to the conclusion, that local but not global comparisons focusing on gender differences in math abilities can be made. Second, our study followed a cross-sectional design. Conclusions about the development of math abilities and gender differences would be more substantive with a longitudinal design of study.

But nevertheless, because of the very early disadvantage of girls in the field of math, some need for action arises. It can be expected that social comparisons and negative experiences at the beginning of girls' school career have negative impact on girls' academic and math self-concept, which in turn has impact on academic interests, later choices of school courses and employment. The existing lack of qualified specialists and employees in the field of math, science and technology leads us to realize that we have to pay attention to early differences in abilities. We have to find its reasons to overcome this phenomenon. Based on the fact that international comparisons revealed different findings regarding gender differences in math, explanations based on biological differences might be less considerable than sociocultural factors. Girls' abilities in math might largely be caused by attitudes towards women and math (Guiso et al., 2008). Thus, this stereotypical view of women and math within the culture and the respective female self-concept build a basis for future research and improvement.

3. Studie II³

Erfassung des akademischen Selbstkonzepts bei Erstklässlern durch eine adaptiv gestaltete Mitschülervergleichsmethode

Zusammenfassung

In der vorliegenden Studie wurde ein neuer, adaptiver Messansatz entwickelt, um das Selbstkonzept von Erstklässlern in den Bereichen Rechnen, Lesen und Schreiben mit dem Fokus auf soziale Vergleichsprozesse möglichst präzise zu erheben. Dazu wurde das Selbstkonzept von 40 Erstklässlern in Einzelsitzungen in den Bereichen Rechnen, Lesen und Schreiben einerseits mit traditionellen Fragebogenitems und andererseits mit einer neu entwickelten Mitschülervergleichsmethode erhoben. Zum Vergleich beider Verfahren wurden Korrelationen mit den Leistungseinschätzungen der Lehrerinnen berechnet. Es zeigt sich, dass die Korrelation der Lehrereinschätzungen mit dem Mitschülervergleichswert sehr hoch ist, während die Korrelation der Lehrereinschätzungen mit dem Fragebogenwert verschwindend gering ausfällt. Diese Unterschiede legen nahe, dass das akademische Selbstkonzept bei Erstklässlern mit der adaptiven Mitschülervergleichsmethode valider erfasst werden kann als mit traditionellen Fragebogenverfahren.

Schlüsselwörter: Selbstkonzept; Erstklässler; Mitschülervergleich; Selbstkonzeptmessung

³ Basierend auf einem zur Veröffentlichung eingereichten Manuskript: Sahr, K., Arndt, D., Opfermann, O. & Leutner, D. (2012, zur Veröffentlichung eingereichtes Manuskript). Erfassung des akademischen Selbstkonzepts bei Erstklässlern durch eine adaptiv gestaltete Mitschülervergleichsmethode.

3.1 Forschungsstand zum akademischen Selbstkonzept

Der Zusammenhang zwischen schulischen Leistungen und dem akademischen Selbstkonzept, also der Vorstellung über die eigenen Fähigkeiten, ist in der über 100-jährigen Selbstkonzeptforschung vielfach belegt (Calsyn & Kenny, 1977; Helmke, 1992; James, 1890; Marsh, 1986; Möller & Köller, 2004; Valentine et al., 2004). Die Wirkrichtung der beiden Faktoren ist dabei jedoch noch nicht eindeutig geklärt (Dickhäuser, 2006; Helmke, 1992). Geht man von einem Einfluss des akademischen Selbstkonzepts auf die Leistung aus, wird dies als Self-Enhancement-Ansatz bezeichnet (Calsyn & Kenny, 1977; Marsh, 1990). Hierbei wird angenommen, dass sich das akademische Selbstkonzept positiv auf das Interesse auswirkt, welches wiederum aufgrund motivationaler Wirkmechanismen die Leistung beeinflusst (Helmke, 1992). Trautwein et al. (2006) konnten diesen Effekt in den PISA-Tests und auch für Schulnoten nachweisen. Köller, Daniels, Schnabel und Baumert (2000) machten anhand von 934 Schülern des Gymnasiums deutlich, dass die Leistungskurswahlen in der 12. Klasse gut durch das Selbstkonzept und das fachspezifische Interesse am Ende der 10. Klasse vorhergesagt werden können. In einer Metaanalyse fassten Valentine et al. (2004) die Ergebnisse von 60 Studien zusammen und konnten ebenso den Effekt eines hohen Selbstkonzepts auf eine bessere Leistung bestätigen. Dem gegenüber stehen Befunde, die zeigen, dass das akademische Selbstkonzept durch die vorherige Leistung bestimmt wird (Skill-Development-Ansatz; Calsyn & Kenny, 1977; Helmke & van Aken, 1995). Vor allem beim viel untersuchten Big-Fish-Little-Pond-Effekt (Gabriel et al., 2010; Marsh, 1987) wird deutlich, dass das akademische Selbstkonzept von den Leistungen bzw. den wahrgenommenen Leistungsrückmeldungen abhängig ist. Hierbei zeigt sich, dass das akademische Selbstkonzept eines Schülers durch die

Einbettung seiner Leistungen in den Klassenkontext gebildet wird. Sowohl der Self-Enhancement- als auch der Skill-Development-Ansatz wurden mehrfach belegt, weshalb beide Wirkrichtungen als gesichert angenommen werden können und man davon ausgeht, dass sie in einem reziproken, dynamischen Zusammenhang zueinander stehen (Marsh & Craven, 2006). Da die zugehörigen Studien sich teilweise stark in den methodischen und auch theoretischen Grundlagen unterscheiden, bleibt bislang allerdings unklar, unter welchen Bedingungen die Leistungen auf das akademische Selbstkonzept wirken oder umgekehrt (Helmke, 1992; Helmke & van Aken, 1995; Valentine et al., 2004).

Der bereits erwähnte Big-Fish-Little-Pond-Effekt zeigt nicht nur den bestehenden Zusammenhang zwischen Leistung und Selbstkonzept, sondern es wird auch deutlich, dass sich Leistungen nicht direkt in einem 1:1-Verhältnis auf das entsprechende Selbstkonzept abbilden lassen. Basierend auf Festingers Theorie der sozialen Vergleiche (1954) resultiert das fachbezogene Selbstkonzept zusätzlich aus den Vergleichen einer Person mit den Leistungen relevanter Bezugsgruppen. Dieser Bezugsgruppeneffekt bewirkt, dass von zwei Personen mit objektiv gleicher Leistung diejenige ein höheres Selbstkonzept besitzt, welche mehr Möglichkeiten zu Abwärtsvergleichen hat, also Teil einer leistungsschwächeren Gruppe ist. Doch auch die sozialen Vergleiche allein sind nicht ausreichend, um die Genese des akademischen Selbstkonzepts zu erklären. Marsh (1986) postulierte in seinem Internal/External-Frame-of-Reference-Modell, dass neben den sozialen (externalen) auch dimensionale (zwischen den einzelnen Fächern einer Person; interne) Vergleiche vorgenommen werden. Da das akademische Selbstkonzept nicht nur global, sondern auch fächerspezifisch ausgeprägt ist (Shavelson et al., 1976), bewirken die Fächervergleiche, dass gute Leistungen in einem Fach das spezifische

Selbstkonzept in einem anderen Fach negativ beeinflussen und vice versa. So zeigte sich auch in der Metaanalyse von Möller et al. (2009), dass das verbale und mathematische Selbstkonzept trotz hoher Zusammenhänge der jeweiligen Leistungen nicht miteinander korrelieren.

Selbstkonzepte im Vor- und Grundschulalter

Kurz nach Veröffentlichung des Selbstkonzeptmodells von Shavelson et al. (1976) wurde die Multidimensionalität des Selbstkonzepts in vielen Studien auch für ältere Kinder und Jugendliche bestätigt (für einen Überblick siehe Marsh, Craven & Debus, 1991). Studien zu den verschiedenen Selbstkonzeptdimensionen bei jüngeren Kindern zwischen fünf und neun Jahren wurden allerdings selten durchgeführt, was hauptsächlich auf zwei Gründe zurückzuführen ist. Schon Shavelson et al. (1976) zeigten, dass die einzelnen Selbstkonzeptdimensionen in diesem Altersbereich noch stark untereinander korrelieren. Erst im Laufe der Entwicklung des Kindes werden die einzelnen Bereiche als distinkte Faktoren deutlich. Die vermeintliche Annahme eines wenig differenzierten Selbstkonzepts dämpfte somit die Forschungstätigkeit im niedrigen Altersbereich. Zusätzlich konnten die üblicherweise verwendeten Paper-Pencil-Fragebögen bei Vor- und Grundschulern aufgrund nicht oder nur wenig vorhandener Lesefähigkeiten kaum eingesetzt werden. Es wurden neue Instrumente zur Selbstkonzepterfassung entwickelt, die aber aufwendiger in ihrer Durchführung waren. So zeigte sich mit Hilfe interviewgeleiteter Befragungen, dass eine Erhebung verschiedener Selbstkonzeptdimensionen auch bei Kindern vom zweiten Schuljahr bis hinunter in den Kindergarten möglich ist (Entwisle et al., 1987; Helmke, 1991; Marsh et al., 1984; Marsh et al., 1991). Marsh et al. (1984, 1991) bestätigten dabei die Annahme, dass die einzelnen Selbstkonzeptdimensionen

bei älteren Kindern besser zu trennen sind als bei jüngeren. Trotzdem konnten alle sieben Einzelbereiche (Aussehen, Beziehung zu Eltern, Beziehung zu Gleichaltrigen, körperliche Fähigkeiten, verbale Fähigkeiten, mathematische Fähigkeiten und akademisches Selbstkonzept) sowie ein globales Selbstkonzept bereits bei Fünfjährigen nachgewiesen werden.

Es ist zu beobachten, dass die Einschätzung der eigenen Fähigkeiten in der Vorschulzeit und den ersten Grundschuljahren stark positiv verzerrt ist und erst im Laufe der Schullaufbahn absinkt. In der Studie von Jacobs et al. (2002) erreichten die Schüler in der ersten Klasse für das mathematische, sprachliche und sportliche Selbstkonzept Werte zwischen 5.2 und 6.3 auf einer siebenstufigen Skala. In den folgenden Schuljahren sanken die Werte in allen Bereichen, wobei das mathematische und verbale Selbstkonzept schneller abnahm als das Selbstkonzept für Sport. Helmke (1991) gab einen Überblick über die aus der Forschung bekannten Gründe für die hohe Einschätzung eigener Fähigkeiten und erwähnte unter anderem, dass die kognitiven Fähigkeiten von Kindern, die jünger als neun Jahre sind, noch nicht hinreichend ausgereift seien, um zwischen ihren Wünschen und der Realität unterscheiden zu können. Andererseits wurde aber auch dargestellt, dass Kinder durchaus realistische Einschätzungen treffen können, nämlich wenn es um die Fähigkeiten anderer Personen geht. Valentine et al. (2004) nannten die noch andauernde kognitive Entwicklung auch als Grund dafür, dass zu Beginn der Grundschule das Selbstkonzept keinen bzw. lediglich einen sehr geringen Einfluss auf die Leistung hat. Der Skill-Development-Ansatz dagegen wurde für die Klassenstufen zwei bis vier von Helmke und van Aken (1995) bestätigt, die den Einfluss von Noten und Leistungstests auf das mathematische Selbstkonzept mit rund 700 Schülern untersuchten. Außerdem zeigte sich, dass Lehrereinschätzungen

in Form von Noten einen größeren Einfluss auf das Selbstkonzept haben als die durch Tests erfasste Fähigkeit.

Eine zusätzliche Erkenntnis aus der Selbstkonzeptforschung der Vor- und Grundschulzeit ergab sich in Bezug auf die Art der Vergleiche, die die Kinder vornehmen, um ihr Fähigkeitsselbstkonzept zu bilden. In der Übersichtsarbeit zum Internal/External-Frame-of-Reference-Modell von Möller und Köller (2004) zeigte sich, dass die Anwendung dimensionaler Vergleiche frühestens ab der vierten Klasse nachgewiesen werden konnte. Dickhäuser und Galfe (2004) fanden weder einen Einfluss dimensionaler noch temporaler (zeitlicher) Vergleiche auf das akademische Selbstkonzept bei Schülern der dritten und vierten Klasse. Hieraus lässt sich ableiten, dass das fachspezifische Selbstkonzept von Schülerinnen und Schülern zu Beginn der Grundschule hauptsächlich auf sozialen Vergleichen beruht.

Erhebung des akademischen Selbstkonzepts im Grundschulalter

Validierte Instrumente zur Erfassung des Selbstkonzepts von Kindern vor dem Schuleintritt und in den ersten Schuljahren sind kaum vorhanden. Der Self-Description-Questionnaire (SDQ; Marsh et al., 1984) ist geeignet, um bereits bei Fünfjährigen sieben verschiedene Selbstkonzeptdimensionen zu erfassen. Allerdings liegt dieser Fragebogen nicht als deutsche Version vor. Lediglich der SDQ-I (Marsh, 1990), der eine Selbstkonzepterfassung bei Kindern zwischen acht und zwölf Jahren möglich macht, wurde kürzlich von Arens et al. (2011) in einer deutschen Version validiert. Die Pictorial Scale von Harter und Pike (1984), die eine einfache und somit kindgerechte Erfassung verschiedener Selbstkonzeptskalen ermöglicht, wurde von Asendorpf und van Aken (1993) in eine deutsche Version übersetzt. Diese Version wurde im Gegensatz zum Original aber erst ab der zweiten Klasse

eingesetzt. Zusätzlich können keine fächerspezifischen Selbstkonzepte differenziert erfasst werden, da nur eine Gesamtskala „Kognitive Kompetenz“ enthalten ist. Ähnlich verhält es sich mit dem Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen (Rauer & Schuck, 2004). Dieser enthält eine allgemeine Skala zum „Selbstkonzept der Schulfähigkeit“. Hierbei werden mit je einem einzelnen Item die Einschätzungen in den Bereichen Mathematik und Lesen abgefragt. Erhebungsinstrumente, die Skalen verschiedener fachspezifischer Selbstkonzepte beinhalten, sind erst ab Ende der Grundschule zur Durchführung angedacht (z.B. SESSKO; Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2002; DISK-Gitter; Rost, Sparfeldt & Schilling, 2007). In den Untersuchungen von Nicholls (1978) sowie auch Helmke (1991) wurde ein anderes Verfahren gewählt. Die Kinder sahen fünf Spielfiguren vor sich, die vom besten bis zum schlechtesten Schüler die Vergleichsgruppe repräsentierten. Nun sollten die Kinder eine eigene Spielfigur innerhalb der Vergleichsgruppe positionieren, je nachdem wo sie ihre Fähigkeiten sahen. So konnte ein direkter Selbstkonzeptwert entnommen werden. Zusammenfassend kann gesagt werden, dass in den meisten Studien mit Grundschulkindern wenige und einfach formulierte Fragen zur Einschätzung akademischer Fähigkeiten Anwendung fanden, denen die Kinder auf mehrstufigen Skalen zustimmen mussten (Dickhäuser & Stiensmeier-Pelster, 2003; Gabriel et al., 2010; Jacobs et al., 2002; Marsh et al., 1991).

Ableitung der Fragestellung

Angesichts der wenigen zur Verfügung stehenden Erhebungsinstrumente und Methoden zur Erfassung des akademischen Selbstkonzepts in der Schuleingangsphase sollten mehr Anstrengungen unternommen werden, diesen Forschungsstand zu erweitern. Zwar zeigte sich beim Selbstkonzept in der Grundschulzeit noch kein Einfluss auf die Leistung (Helmke & van Aken, 1995), aber spätestens beim Übergang in die Sekundarstufe stellt ein hohes akademisches Selbstkonzept einen wichtigen Prädiktor der positiven Bewältigung dieser Veränderungssituation dar (für einen Überblick siehe Valentine et al., 2004). Da Grundschüler ihr akademisches Selbstkonzept anhand der Lehrereinschätzung ihrer Fähigkeiten bilden (Dickhäuser & Meyer, 2006; Helmke & van Aken, 1995), sollte es auch schon zu Beginn der Schullaufbahn möglich sein, das akademische Selbstkonzept fächerspezifisch und valide zu erheben.

Weiterhin resultierten die bisher erhobenen Selbstkonzepte junger Kinder in sehr hohen Werten auf den jeweiligen Skalen (z.B. Jacobs et al., 2002). In den meisten Studien sollen die Kinder entweder globale Aussagen treffen (z.B. „Wie gut bist du im Lesen?“) oder Vergleiche mit einer großen Gruppe von Mitschülern bzw. Gleichaltrigen durchführen (z.B. „Wie gut kannst du rechnen im Vergleich zu deinen Mitschülern?“). Dies gilt auch für die etwas andere Erhebungsmethode bei Nicholls (1978) und Helmke (1991), denn obwohl die Kinder eine direkte Einschätzung angeben konnten, mussten sie diese Einschätzung erst einmal anhand einer wenig definierten Vergleichsgruppe bilden (allgemeiner Vergleich zu anderen Kindern). Aufgrund des in diesem Altersbereich noch eingeschränkten kognitiven Entwicklungsstandes könnten mit den bisherigen Erhebungsmethoden die hohen Selbstkonzeptwerte zustande kommen, die aber nicht die wahre Einschätzung ihrer

Fähigkeiten abbildet, sondern eher einen Wunsch darstellt.

Somit ist eine Methode nötig, die zwar einerseits die sozialen Vergleiche berücksichtigt, andererseits aber nur wenige Informationseinheiten zur gleichen Zeit benötigt, um eine kognitive Überforderung der Kinder zu vermeiden. Ist dies gelungen, stehen die neuen Selbstkonzeptwerte in einem größeren Zusammenhang zu den Fähigkeiten als die Selbstkonzeptwerte der traditionellen Fragebogenverfahren.

3.2 Methode

Stichprobe

An der Untersuchung nahmen 40 Kinder aus zwei Klassen einer Grundschule in Nordrhein-Westfalen teil. Das durchschnittliche Alter betrug 79.4 Monate ($SD = 4.7$). Zum Zeitpunkt der Erhebung befanden sich die Kinder seit vier Monaten in der ersten Klasse. Das Geschlechterverhältnis kann mit 19 Jungen und 21 Mädchen als ausgeglichen bezeichnet werden.

Material und Vorgehensweise

Ziel der Untersuchung war die Erhebung des domänenspezifischen Selbstkonzepts in den Bereichen Rechnen, Lesen und Schreiben. Hierzu wurden die Kinder einzeln in einem ruhigen Raum innerhalb des Schulgebäudes von einem Testleiter befragt. Nach einer kurzen Einleitung erfolgte die Erfassung des Selbstkonzepts mit einem herkömmlichen Fragebogenverfahren. Die Kinder erhielten nacheinander zwei Beispielitems (Ich bin gut in der Schule; Ich kann gut malen), mit denen das Antwortformat eingeübt wurde und sie sich an die Situation gewöhnen konnten. Daraufhin wurden den Kindern zu jedem Leistungsbereich je vier Items präsentiert. Drei Items stammten aus der Studie von Dickhäuser und Stiensmeier-Pelster (2003;

Ich bin gut im Rechnen, Rechnen fällt mir leicht, Ich bin besser im Rechnen als die meisten Kinder) und wurden um ein Item ergänzt, welches zusätzlich die dimensionale Komponente erfasste (Ich bin besser im Rechnen als in den anderen Fächern). Die Einschätzung dieser Items erfolgte auf einer 5-stufigen Likert-Skala, die einerseits mündlich durch den Testleiter (von „stimmt gar nicht“ [1] bis „stimmt ganz genau“ [5]) und andererseits über fünf Smileys (sehr traurig bis sehr fröhlich) präsentiert wurde. Die Zuordnung der mündlichen Einschätzung zu dem entsprechenden Smiley erfolgte durch das gleichzeitige Zeigen des Smileys und der Nennung der jeweiligen Zustimmungsmöglichkeit durch den Testleiter während der zwei Beispielitems. Die Erstklässler konnten ihre Antwort mündlich geben und/oder auf den entsprechenden Smiley zeigen.

Im Anschluss erfolgte die Befragung der Kinder mit der neu entwickelten, adaptiven Mitschülervergleichsmethode. Abbildung 3.1 zeigt zur Verdeutlichung ein fiktives Beispiel einer Klasse. Hierzu wurden im Vorfeld der Untersuchung von den zwei Lehrerinnen Rangfolgen der Kinder ihrer Klassen in den Bereichen Rechnen, Lesen und Schreiben gebildet. Eine Vorgabe über eine konkrete Anzahl an Rängen wurde nicht gegeben. Eine Zuordnung mehrerer Schüler zu einem Rang war möglich. Im fiktiven Beispiel (Abbildung 3.1) zeigt sich dies, indem pro Rang zwei Kinder von der Lehrerin verortet wurden. So vergaben die Lehrerinnen in dieser Studie zwischen acht und zehn Rangplätze pro Leistungsbereich. Anhand dieser Ränge wurden für jedes Kind in jedem Bereich fünf Mitschüler ausgewählt, von denen jeweils zwei als besser, zwei als schlechter und ein Mitschüler als ungefähr gleich gut durch die Lehrerin eingeschätzt wurden. Die Auswahl erfolgte adaptiv, so dass die fünf Vergleichsmitschüler in der Nähe des Rangplatzes des Kindes verortet waren.

Im fiktiven Beispiel (Abbildung 3.1) wird dies z.B. an der Schülerin Laura deutlich. Für sie wurden zwei gemäß Lehrereinschätzung bessere Mitschüler (Sarah, Rang Eins & Paul, Rang Zwei), ein Mitschüler, der die gleiche Leistung zeigt (Leon, Rang Drei) und zwei schlechtere Mitschüler (Anna, Rang Vier & Ben, Rang Fünf) ausgewählt. War diese Auswahl aufgrund eines sehr niedrigen oder sehr hohen Rangplatzes des Kindes nicht möglich, so wurden trotzdem fünf Mitschüler ausgewählt, indem die nicht möglichen besseren oder schlechteren Vergleichsmitschüler durch nahe liegende andere Mitschüler ersetzt wurden. Deutlich wird dies im fiktiven Beispiel an den Schülern Ben und Lilly, für die keine Vergleiche mit schlechteren Schülern möglich sind. Sie erhielten demzufolge zwei Vergleiche mit Mitschülern, die ähnliche bzw. nur leicht bessere Leistungen aufwiesen. Jedes Kind wurde also 15 Mal (3 Bereiche x 5 Mitschüler) gefragt: „Wer ist besser? Du oder [*Name des Mitschülers*]?“ Die Befragung erfolgte als Forced-Choice-Verfahren, was bedeutet, dass die Kinder entweder sich selbst oder den jeweiligen Mitschüler als besser benennen sollten. Konnte das Kind trotz Nachfrage keinen von beiden als besser auswählen, wurde alternativ ein anderer Mitschüler mit vergleichbarem Rang zum Vergleich vorgegeben. Aus ökonomischen und motivationalen Gründen erfolgte kein vollständiger Paarvergleich (Vergleich mit allen Mitschülern).

Zusätzlich bearbeiteten alle Kinder in einer Einzelsitzung einen Test zur Erfassung der mathematischen Fähigkeiten, der auf dem Modell der mathematischen Entwicklung von Fritz und Ricken (2008) basiert.

Name	Mitschüler	Sarah	Tim	Marie	Paul	Laura	Leon	Anna	Luca	Lilly	Ben
Kind	Rang	1	1	2	2	3	3	4	4	5	5
Sarah	1		X	X			X		X		X
Tim	1	X			X	X		X		X	
Marie	2	X	X		X			X			X
Paul	2	X	X	X					X	X	
Laura	3	X			X		X	X			X
Leon	3		X	X		X			X	X	
Anna	4		X	X					X	X	X
Luca	4	X			X			X		X	X
Lilly	5			X		X		X	X		X
Ben	5				X		X	X	X	X	

Abbildung 3.1: Beispielhafte Darstellung der Rangfolgen innerhalb einer Klasse und der adaptiven Auswahl der Vergleichsmitschüler pro Kind

Auswertung der Daten

Da aufgrund der theoretischen Überlegungen der soziale Vergleich als hauptursächlich für die Selbstkonzeptgenese im Anfangsunterricht angesehen wird, wurde dies auch in der Mitschülervergleichsmethode fokussiert. Um das traditionelle Fragebogenverfahren mit der neuen Methode vergleichen zu können, wurde für die Auswertung nur das Item des sozialen Vergleichs betrachtet („Ich bin besser im Rechnen/Lesen/Schreiben als die meisten Kinder“).

Für die Auswertung des Mitschülervergleichs wurde auf Basis der Antworten für jedes Kind ein – gegenüber dem Lehrer-Rating – neuer Rangplatz gebildet, welcher den aus den Vergleichen mit den Mitschülern resultierenden Selbstkonzeptwert repräsentiert. Dazu wurde für jedes Kind die Anzahl positiver Vergleiche mit den fünf Mitschülern ausgezählt (Wie oft hat das Kind gesagt, dass es besser ist als der jeweilige Mitschüler?). Diese Anzahl wurde – ausgehend vom jeweiligen schwächsten Vergleichsmitschüler und der Rangfolge der Lehrereinschätzung nach oben folgend – abgezählt. Das Kind erhielt den Rangplatz desjenigen Mitschülers zugeordnet, bei dem das Abzählen endete. Abbildung 3.2 zeigt die fiktiven Antworten der Kinder und die dazugehörigen neuen Rangplätze. Am Beispiel von Luca lässt sich die Vorgehensweise verdeutlichen. Luca sollte sich mit Sarah, Paul, Anna, Lilly und Ben vergleichen und hat dreimal geantwortet, dass er besser sei als der jeweilige Mitschüler (nämlich im Vergleich mit Sarah, Anna und Ben). Diese drei positiven Vergleiche werden nun vom schwächsten Vergleichsmitschüler (Ben) über den zweitschwächsten Mitschüler (Lilly) bis zum drittschwächsten Mitschüler (Anna) abgezählt. Luca positioniert sich somit auf dem Rang von Anna und damit auf Rang vier der ursprünglichen Lehrereinschätzung. Aufgrund der vor der Erhebung adaptiv ausgewählten Mitschüler konnten die

Kinder nun einen besseren, gleichen oder schlechteren Rangplatz erhalten als durch die Lehrerin eingeschätzt. Nicht mögliche Vergleiche mit schlechteren Mitschülern (z.B. im Fall von Ben oder Lilly) stellen insofern kein Problem dar, als aufgrund des niedrigen Leistungsstandes sowieso ein niedriges Selbstkonzept erwartet wird. Sollten Ben oder Lilly z.B. ein höheres Selbstkonzept aufweisen als aufgrund ihrer Leistung zu erwarten ist, würde sich dies darin zeigen, dass sie sich auch bei den besseren Vergleichsmitschülern öfter als besser einschätzen. Sehen sie ihre Mitschüler dagegen immer als besser an, so verbleiben sie erwartungskonform mit ihrem Selbstkonzept auf dem letzten Rang.

Um zu überprüfen, ob die adaptive Vorgehensweise nicht a priori hohe Zusammenhänge mit der Leistungseinschätzung durch die Lehrerin hervorruft, wurden zusätzlich für jedes Kind fünf Zufallsränge gebildet, indem dem Kind durch Würfeln fünf Ränge aus den fünf ausgewählten Mitschülern zugeteilt wurden. So erhielt das Kind z.B. bei einer gewürfelten Eins den Rangplatz des ersten Vergleichsmitschülers als neuen Selbstkonzeptwert. Im fiktiven Beispiel (Abbildung 3.2) würde Marie bei einer gewürfelten Eins den Rangplatz von Sarah erhalten, bei einer gewürfelten Zwei den Rangplatz von Tim, bei einer gewürfelten Drei den Rangplatz von Paul usw.

In dieser Studie wurden drei Merkmale (Rechnen, Lesen und Schreiben) jeweils mit drei Methoden (Lehrereinschätzung, Fragebogenitem und Mitschülervergleichsmethode) untersucht, deren Korrelationen sich – zur Bestimmung der konvergenten bzw. diskriminanten Validität der Messungen – in einer Multitrait-Multimethod-Matrix anordnen lassen (Campbell & Fiske, 1959). Aufgrund vorhandener Rangplatzbindungen wurden γ -Korrelationen als Maß des schwach monotonen Zusammenhangs berechnet.

Name	Mitschüler	Sarah	Tim	Marie	Paul	Laura	Leon	Anna	Luca	Lilly	Ben	neuer
Kind	Rang	1	1	2	2	3	3	4	4	5	5	Rang
Sarah	1		0	1			1		1		1	2
Tim	1	1			1	1		1		1		1
Marie	2	0	1		1			1			1	1
Paul	2	0	1	0					1	1		2
Laura	3	0			1		0	0			1	4
Leon	3		0	0		1			1	1		3
Anna	4		0	1					1	1	1	2
Luca	4	1			0			1		0	1	4
Lilly	5			1		1		0	1		1	3
Ben	5				0		0	1	0	0		5

Abbildung 3.2: Beispielhafte Darstellung der Antworten der Kinder (0 – Kind schätzt Mitschüler als besser ein; 1 – Kind schätzt sich als besser ein) sowie der sich daraus neu ergebenden Rangzuordnungen.

3.3 Ergebnisse

Die Korrelationen der erhobenen Variablen sind in Tabelle 3.1 dargestellt, wobei die Fragebogen- und Testwerte der einfacheren Darstellbarkeit wegen vor Berechnung der Korrelationen invertiert wurden. Die Korrelationen in den Heterotrait-Monomethod-Dreiecksmatrizen (z.B. Lehrereinschätzung Lesen und Lehrereinschätzung Schreiben; kursiv gesetzte Werte in Tabelle 3.1) sind mit $.283 \leq .748$ alle statistisch signifikant. Dies weist darauf hin, dass die verschiedenen Traits (Rechnen, Lesen, Schreiben) unabhängig davon, mit welcher Methode (Lehrereinschätzung, Fragebogen, Mitschülervergleich) sie gemessen wurden, nicht klar voneinander diskriminiert werden. Für die Bewertung der konvergenten Validität sind die Monotrait-Heteromethod-Diagonalen von Bedeutung (z.B. Lehrereinschätzung Schreiben und Fragebogen Schreiben; fett gesetzte Werte in Tabelle 3.1). Die Korrelation von Fragebogen und Lehrereinschätzung ist nur beim Lesen signifikant, $\gamma = .297, p = .031$. Dasselbe gilt für die Korrelation von Fragebogen und Mitschülervergleich, $\gamma = .331, p = .049$. Angesichts der vergleichsweise niedrigen Korrelationen kann dies allerdings nur als schwacher Hinweis auf konvergente Validität der Fragebogenwerte beim Lesen gewertet werden. Beim Rechnen und Schreiben korrelieren die Fragebogenwerte dagegen sehr schwach und statistisch nicht signifikant mit der Lehrereinschätzung und den Mitschülervergleichswerten, was als Hinweis auf mangelnde konvergente Validität der Fragebogenwerte anzusehen ist. Anders stellt sich dagegen der Zusammenhang von Mitschülervergleich und Lehrereinschätzung dar, für den sich mittlere bis hohe und statistisch signifikante Korrelationen ergeben; $\gamma_{\text{Rechnen}} = .715, p < .001, \gamma_{\text{Lesen}} = .811, p < .001, \gamma_{\text{Schreiben}} = .504, p < .001$, was als Hinweis auf gute bis sehr gute konvergente Validität der Mitschülervergleichswerte anzusehen ist.

Darüber hinaus bestätigten Korrelationen mit dem drei Monate zuvor durchgeführten Mathematiktest diese Ergebnisse zur konvergenten Validität. Der Testwert korreliert mit dem Lehrereinschätzung Rechnen in Höhe von $\gamma = .446$, $p < .001$, und mit den Mitschülervergleich Rechnen in Höhe von $\gamma = .353$, $p = .011$. Mit dem Fragebogenwert Rechnen korreliert der Testwert mit $\gamma = .122$ dagegen nur sehr niedrig und nicht signifikant. Bemerkenswert sind allerdings die Heterotrait-Heteromethod-Korrelationen des Mathematiktests, also die Korrelationen mit Messwerten, bei denen weder die Methode noch der Trait übereinstimmen. So korreliert der Mathematiktest z.B. mit dem Lehrereinschätzung Lesen in gleicher Höhe ($\gamma = .430$) wie mit dem Lehrerteil Rechnen ($\gamma = .446$). Dies deckt sich mit dem zuvor berichteten Befund, dass Rechnen und Lesen im Lehrerteil nicht hinreichend diskriminiert werden ($\gamma = .748$). Ob die hohen Korrelationen Urteilsfehler im Sinne einer impliziten Persönlichkeitstheorie darstellen oder tatsächlichen Leistungskorrelationen entsprechen, lässt sich in der vorliegenden Studie nicht feststellen, da die untersuchten Kinder keinen Lesetest bearbeitet haben.

Interessant sind auch die Monotrait-Heteromethod- und die Heterotrait-Heteromethod-Korrelation des Mathematiktests mit den anderen Methoden. Für konvergente Validität des Mitschülervergleichs spricht die hohe Korrelation von Mathematiktest und Mitschülervergleich ($\gamma = .353$), die vergleichbar hoch liegt wie die Korrelation von Mathematiktest und Lehrereinschätzung. Dagegen korrelieren Mathematiktest und Fragebogen nur schwach und nicht signifikant ($\gamma = .122$). Letzteres spricht gegen konvergente Validität des Fragebogens. Mathematiktest und Fragebogen Lesen ($\gamma = .307$) sowie Mathematiktest und Mitschülervergleich Lesen ($\gamma = .495$) korrelieren vergleichsweise hoch. Die hohe Korrelation von Mathematiktest und Mitschülervergleich Lesen entspricht der Korrelation mit dem

Lehrereinschätzung Lesen ($\gamma = .430$) und unterstützt die zuvor angesprochene mangelnde Diskriminierbarkeit der Mathematik- und Leseleistung. Die ebenfalls hohe Korrelation des Mathematiktests mit dem Fragebogen Lesen ($\gamma = .307$) steht im Einklang mit den nicht allzu hohen, aber statistisch signifikanten Korrelationen des Fragebogen Lesens mit der Lehrereinschätzung Lesen ($\gamma = .297$) und dem Mitschülervergleich Lesen ($\gamma = .331$) und unterstützt – die Überlegungen zur Diskriminierbarkeit von Lese- und Rechenleistungen mit einschließend – den zuvor berichteten schwachen Hinweis auf konvergente Validität zumindest der Fragebogenwerte Lesen.

Abschließend wurde untersucht, ob die Mitschülervergleichsmethode möglicherweise zu Werten führt, deren Korrelation mit der Lehrereinschätzung nicht zuletzt schon deswegen hoch ausfallen sollte, weil die Mitschüler, mit denen die Kinder sich zu vergleichen haben, an der Lehrereinschätzung orientiert ausgewählt werden (was letztendlich die Adaptivität der Methode kennzeichnet). Zu diesem Zweck wurden Korrelationen der Lehrereinschätzung mit jeweils fünf Zufallswerten aus fiktiven Mitschülervergleichen berechnet. Es ergaben sich γ -Werte für das Rechnen zwischen $.069$ ($p \geq .05$) und $.541$ ($p < .001$), für das Lesen zwischen $.249$ ($p \geq .05$) und $.631$ ($p < .001$) und für das Schreiben zwischen $.032$ ($p \geq .05$) und $.559$ ($p < .001$). Beim Rechnen und Lesen erreichte allerdings keine Korrelation der Zufallswerte das Niveau der Originaldaten. Auch beim Schreiben zeigte sich, dass keine Korrelation signifikant höher ausfiel als die Korrelation der Originaldaten. Im Übrigen ergaben sich nur dann und in Einzelfällen hohe Korrelationen zwischen Lehrereinschätzung und zufällig generiertem Mitschülervergleich, wenn die zufällig generierten Mitschülervergleiche sowieso hoch mit den Originalwerten der Mitschülervergleichen übereinstimmten.

Tabelle 3.1: Multitrait-Multimethod-Matrix der Selbstkonzeptmessung (Rechnen, Lesen und Schreiben als Traits; Lehrereinschätzung, Fragebogen und Mitschülervergleich als Methods; ergänzt um Korrelationen mit einem Mathematiktest)

		Lehrereinschätzung			Fragebogen			Mitschülervergleich		
		Rechnen	Lesen	Schreiben	Rechnen	Lesen	Schreiben	Rechnen	Lesen	Schreiben
Lehrer- einschätzung	Rechnen									
	Lesen	<i>.748***</i>								
	Schreiben	<i>.283*</i>	<i>.377**</i>							
Fragebogen	Rechnen	-.041	-.055	.049						
	Lesen	.238	.297*	.262	<i>.523***</i>					
	Schreiben	-.026	.011	.013	<i>.549**</i>	<i>.461**</i>				
Mitschüler- vergleich	Rechnen	.715***	.595***	.307*	.080	.369*	-.050			
	Lesen	.583***	.811***	.432***	.151	.331*	.105	<i>.474**</i>		
	Schreiben	.357**	.438**	.504***	.161	.202	.233	<i>.478***</i>	<i>.517***</i>	
Test	Rechnen	.446***	.430***	.573***	.122	.307*	-.124	.353*	.495***	.256

Anmerkungen. * $p < .05$, ** $p < .01$, *** $p < .001$; konvergente Korrelationen (heterotrait-monomethod) kursiv, diskriminante Korrelationen (monotrait-heteromethod) fett hervorgehoben; Fragebogen- und Testwert invertiert.

3.4 Diskussion

Ziel der vorliegenden Studie war es ein neues Verfahren vorzustellen und zu validieren, das – unter Berücksichtigung der geringen Lesekenntnisse und der noch nicht voll ausgereiften kognitiven Fähigkeiten der untersuchten Altersgruppe – eine bessere Erhebung des akademischen Selbstkonzepts zu Beginn der Grundschule erlauben sollte. Die Einordnung innerhalb einer Rangreihe von Vergleichskindern, wie sie von Nicholls (1978) und Helmke (1991) verwendet wurde, bot einen vielversprechenden Ansatz, um einen direkten Selbstkonzeptwert von jungen Kindern zu erhalten. Allerdings ist der Vergleich zu einer großen und undefinierten Gruppe von Kindern aufgrund des kognitiven Entwicklungsstandes als fehleranfällig anzusehen. So wurde der adaptive Mitschülervergleich entwickelt, bei dem die Kinder nur Einzelvergleiche zu bekannten Mitschülern vornehmen mussten und sich aus ihren Antworten ein Selbstkonzeptwert über die Einordnung in die Rangreihe von Mitschülern ergab. Die berichteten Ergebnisse bestätigen die Brauchbarkeit des neuen Verfahrens. Im Rahmen der konvergenten und diskriminanten Validierung des Verfahrens zeigen sich hohe Korrelationen mit den Leistungseinschätzungen der Lehrerinnen, was als Hinweis auf konvergente Validität der Mitschülervergleichsmethode zu werten ist. Parallel zum Mitschülervergleich wurde ein Fragebogen eingesetzt, wie er typischerweise in Studien zum akademischen Selbstkonzept verwendet wird (Dickhäuser & Stiensmeier-Pelster, 2003; Gabriel et al., 2010). Für das für die Analysen ausgewählte und den sozialen Vergleich repräsentierende Item konnte bezüglich des Selbstkonzepts Lesen nur schwache, bezüglich der Selbstkonzepte Rechnen und Schreiben dagegen keine Hinweise auf konvergente Validität gefunden werden. Hinsichtlich diskriminanter Validität zeigte sich, dass sich die drei Selbstkonzepte nicht klar voneinander trennen lassen, und

zwar unabhängig davon, mit welchem Verfahren sie erfasst werden.

Diese Befunde weisen darauf hin, dass die im Bereich der Selbstkonzeptforschung typischerweise angewendeten Fragebogenverfahren am Anfang der Grundschule möglicherweise etwas anderes abbilden als der direkte Mitschülervergleich. Da das Selbstkonzept bei Erstklässlern anhand von Leistungsrückmeldungen gebildet wird und sie diese Rückmeldungen in dieser Zeit vorrangig von den jeweiligen Lehrern erhalten (Helmke & van Aken, 1995), sollte das Selbstkonzept mit der Leistungseinschätzung der Lehrer eng zusammenhängen. Dies war beim Rechnen und Schreiben der Fall, wenn das Selbstkonzept per Mitschülervergleich, nicht aber wenn das Selbstkonzept per Fragebogen erhoben worden war. Beim Lesen ergab sich beim Vergleich von Lehrereinschätzung und Selbstkonzept eine sehr hohe Korrelation, sofern der Mitschülervergleich betrachtet wird, und eine schwache, aber statistisch signifikante Korrelation, wenn der Fragebogen herangezogen wird. Die Sonderstellung des Lesens könnte darauf zurückgeführt werden, dass die Kinder vermehrt auch im außerschulischen Bereich lesen und Rückmeldungen über ihre Lesefähigkeit öfter erhalten als Rückmeldungen im Rechnen oder Schreiben. Somit sind sie geübter, ihre Lesekompetenz auch global im Vergleich zu anderen Kindern einzuschätzen.

Als Kritikpunkt am Mitschülervergleich könnte angeführt werden, dass die hohen Korrelationen der Ergebnisse mit der Lehrereinschätzung durch die individuell adaptierte Auswahl derjenigen Mitschüler bedingt sein könnten, mit denen sich die Kinder zu vergleichen hatten. Dies konnte anhand der Analyse von artifiziellen Werten, die den Kindern bei gegebener Mitschüleraushwahl zufällig zugeordnet worden waren, als wenig wahrscheinlich dargelegt werden. In den betrachteten Leistungsbereichen ergaben sich beim Vergleich der

Mitschülervergleichswerte mit der Lehrereinschätzung oft Nullkorrelationen und die meisten Korrelationen waren niedriger als die jeweiligen Korrelationen auf Basis der Originaldaten. In jedem Fall aber war keine Korrelation auf Basis der Zufallsdaten statistisch signifikant höher als die jeweilige Korrelation auf Basis der Originaldaten.

Ein Nachteil der Mitschülervergleichsmethode ist, dass z.B. im Gegensatz zu Helmke (1991) keine absoluten Werte generiert werden, mit denen Klassen oder einzelne Schüler über Klassen hinweg miteinander verglichen werden können. Jeder Mitschülervergleichswert kann nur im Kontext der jeweiligen Klasse interpretiert werden. Dies stellt allerdings kein Problem dar, denn die empirische Forschung kann dieses Verfahren z.B. in Form von Korrelationsstudien, wie dieser vorliegenden Arbeit, anwenden. Auch in der Praxis kann der Mitschülervergleich innerhalb der einzelnen Klassen Anwendung finden. Lehrer sollten mithilfe dieses Verfahrens erkennen können, wie das akademische Selbstkonzept eines Schülers ausgeprägt ist und sollten gegebenenfalls durch angepasste Rückmeldungen korrigierend auf negative Selbstkonzepte einwirken. Die bisher noch vergleichsweise aufwendige Vor- und Nachbereitung der Datenerhebung könnte zudem durch ein Computerprogramm, welches z.B. die Auswahl der Vergleichsmitschüler vorgibt und die Mitschülervergleichswerte berechnet, erleichtert werden.

Abschließend ist festzustellen, dass sich der Mitschülervergleich im Anfangsunterricht als überlegen zumindest gegenüber den bisher üblicherweise zum Einsatz kommenden Fragebogenitems des sozialen Vergleichs herausgestellt hat. Ob dies auch in der weiteren Schullaufbahn zutrifft, könnte eine Fragestellung zukünftiger Untersuchungen sein.

4. Studie III⁴

Unterschiede im akademischen Selbstkonzept zwischen Mädchen und Jungen in der 1. Klasse

Zusammenfassung

In dieser Arbeit wurde untersucht, inwieweit sich Mädchen und Jungen bereits zu Beginn der Schulzeit in ihren akademischen Selbstkonzepten unterscheiden. Dazu wurden 142 Erstklässler mit zwei unterschiedlichen Methoden zu ihren Fähigkeitseinschätzungen in den Bereichen Rechnen, Lesen und Schreiben befragt. Es zeigte sich lediglich für das Rechnen ein Geschlechtsunterschied, wenn die Kinder mit herkömmlichen Fragebogenitems nach ihren Fähigkeiten gefragt wurden. Hierbei gaben Mädchen signifikant höhere Einschätzungen ihrer rechnerischen Fähigkeiten an als Jungen. Zur Erklärung dieses Unterschieds wurden mögliche Einflussfaktoren wie Leistungsunterschiede und geschlechtsspezifische Fähigkeitseinschätzungen der Lehrer betrachtet. Allerdings konnten dabei keine signifikanten Geschlechtsunterschiede nachgewiesen werden, wodurch Leistung und Lehrereinschätzung als Ursachen des Selbstkonzeptunterschieds ausgeschlossen werden können.

Schlüsselwörter: Akademisches Selbstkonzept, Geschlechtsunterschiede, Erstklässler

⁴ Basierend auf einem unveröffentlichtem Manuskript: Sahr, K., Arndt, D., Opfermann, O. & Leutner, D. (2012, unveröffentlichtes Manuskript). Unterschiede im akademischen Selbstkonzept zwischen Mädchen und Jungen in der 1. Klasse.

4.1 Einleitung

Gute mathematische und verbale Fähigkeiten sind grundlegende Voraussetzungen für eine erfolgreiche Schullaufbahn und späteren beruflichen Erfolg. Ein mehrfach empirisch belegter Zusammenhang findet sich zwischen akademischen Leistungen und den entsprechenden Selbstkonzepten (Calsyn & Kenny, 1977; Helmke, 1992; Marsh, 1986; Möller & Köller, 2004; Valentine et al., 2004). Hierbei konnten zwei Wirkrichtungen nachgewiesen werden. Zum einen gilt der Self-Enhancement-Ansatz, der besagt, dass das akademische Selbstkonzept einen Einfluss auf die akademischen Leistungen hat (Calsyn & Kenny, 1977; Marsh, 1990). Dies bedeutet, dass eine Person, die davon überzeugt ist, gute akademische Fähigkeiten zu besitzen – also ein höheres akademisches Selbstkonzept aufweist, bei gleichen kognitiven Voraussetzungen in Zukunft bessere Leistungen erbringen wird als eine Person, die weniger stark an ihre akademischen Fähigkeiten glaubt – also ein niedrigeres Selbstkonzept besitzt. Im Gegensatz dazu konnten Studien auch den Skill-Development-Ansatz nachweisen, der besagt, dass sich die akademische Leistung auf das Selbstkonzept auswirkt (Calsyn & Kenny, 1977; Helmke & van Aken, 1995). Deutlich wurde dieser Ansatz vor allem in den Arbeiten zum Big-Fish-Little-Pond-Effekt (Marsh, 1987), da hierbei das akademische Selbstkonzept eines Schülers bzw. einer Schülerin durch das Verhältnis seiner bzw. ihrer Leistungen zur durchschnittlichen Leistung in der Klassen beeinflusst wird. Beide Ansätze wurden mehrfach empirisch bestätigt, so dass sowohl der Self-Enhancement- als auch der Skill-Development-Ansatz derzeit als gültig angesehen werden können. Allerdings ist bisher noch nicht hinreichend geklärt, in welchen Situationen die Leistungen auf das akademische Selbstkonzept wirken oder wann das akademische Selbstkonzept einen Einfluss auf die Leistungen hat (Helmke, 1992; Helmke & van Aken, 1995;

Valentine et al., 2004). Ein vielversprechender Ansatz scheint das Reciprocal-Effects-Model (Marsh & Craven, 2006) zu sein, welches einen reziproken, dynamischen Zusammenhang zwischen akademischen Leistungen und dem entsprechenden Selbstkonzept annimmt.

Selbstkonzept im Vor- und Grundschulalter

Shavelson et al. (1976) beschreiben eine hierarchische Struktur des Selbstkonzepts, was bedeutet, dass das allgemeine Selbstkonzept sich aus dem akademischen und nicht-akademischen Selbstkonzept zusammensetzt. Das akademische Selbstkonzept teilt sich wiederum in die Selbstkonzepte der einzelnen Schulfächer, während das emotionale, physische und soziale Selbstkonzept Teile des nicht-akademischen Selbstkonzepts sind. Hierbei kann eine weitere Aufteilung in Subdimensionen (z.B. Selbstkonzept bezüglich Freundschaften) vorgenommen werden. Die Existenz vieler unterschiedlicher Selbstkonzeptdimensionen wird als Multidimensionalität bezeichnet. Für Kinder im Vor- und Grundschulalter sei diese Struktur des Selbstkonzepts allerdings noch nicht anwendbar, weil sie ihre Erfahrungen noch nicht adäquat und strukturiert in ihr Selbstkonzept integrieren. Erst mit steigendem Alter und verbesserten verbalen Fähigkeiten differenzieren sich die verschiedenen Selbstkonzeptdimensionen heraus. Untersuchungen mit Vor- und Grundschulern bestätigten die Annahme von Shavelson et al. (1976) aber nur insofern, als die Korrelationen zwischen den einzelnen Selbstkonzeptdimensionen sehr hoch sind und mit dem Alter geringer werden (Asendorpf & van Aken, 1993; Marsh et al., 1991). Allerdings konnten die sieben Dimensionen des Self-Description-Questionnaire (SDQ: Aussehen, Beziehung zu Eltern, Beziehung zu Gleichaltrigen, körperliche Fähigkeiten, verbale Fähigkeiten, mathematische Fähigkeiten und

darüber hinaus das akademische Selbstkonzept) sowie auch ein globales Selbstkonzept bereits bei Fünfjährigen faktorenanalytisch nachgewiesen werden (Marsh et al., 1991).

Übereinstimmend zeigten die Forschungsbefunde eine überdurchschnittlich positive Selbstkonzepteinschätzung bei jüngeren Kindern in allen Dimensionen des Selbstkonzepts (Eccles et al., 1993; Jacobs et al., 2002; Martschinke & Kammermeyer, 2006; Nicholls, 1978; Poloczek et al., 2011). Während in der KILIA-Studie das mathematische und verbale Selbstkonzept von der ersten zur zweiten Klasse konstant hoch bleiben (Martschinke & Kammermeyer, 2006), berichtet Helmke (1998) in der Logik-Studie von einem starken Einschnitt des Selbstkonzepts in Mathematik und Deutsch für den gleichen Zeitraum. Bei längsschnittlichen Untersuchungen konnte aber generell ein Abfall der unterschiedlichen Selbstkonzeptdimensionen von Beginn der Schulzeit bis ins Erwachsenenalter festgestellt werden (Helmke, 1998; Jacobs et al., 2002). Als Gründe der überdurchschnittlich positiven Selbstkonzepte bei jungen Kindern führte Nicholls (1978) ähnlich wie Shavelson et al. (1976) die noch andauernde kognitive Entwicklung an. Erst ab einem Alter von neun Jahren seien die Kinder in der Lage, unterschiedliche Informationsquellen bezüglich ihrer Leistung zu nutzen und in ein adäquates Selbstkonzept zu übertragen. Weiterhin können jüngere Kinder schlecht zwischen ihren Wünschen und der Realität unterscheiden (Stipek & Hoffmann, 1980) und geben somit eher ihre gewünschte Einschätzung als tatsächliche Gegebenheiten an. Aufgrund dessen verringern sich die optimistischen Selbstkonzepte im Laufe der Schulzeit und erreichen zum Ende der Pubertät ein realistisches Niveau (Jacobs et al., 2002).

Die Debatte über den kausalen Zusammenhang zwischen Leistung und den

entsprechenden akademischen Selbstkonzeptdimensionen wird auch für den Grundschulbereich geführt. Valentine et al. (2004) nennen die noch unausgereifte kognitive Entwicklung als Grund dafür, dass das Selbstkonzept zu Beginn der Schulzeit nur einen sehr geringen Einfluss auf Leistungsparameter hat. Helmke und van Aken (1995) konnten zeigen, dass von der zweiten bis zur vierten Klasse hauptsächlich der Skill-Development-Ansatz wirkt, wobei Noten das Selbstkonzept mehr beeinflussen als die mit wissenschaftlichen Tests erhobenen Leistungen. In ihrem Übersichtsartikel kommen auch Marsh und Craven (2006) zu dem Schluss, dass am Anfang der Schulzeit der Skill-Development-Ansatz stärker zum Tragen kommt als der Self-Enhancement-Ansatz. Allerdings konnte schon ab der zweiten Klasse ein reziproker Zusammenhang im Sinne des Reciprocal-Effects-Model (Marsh & Craven, 2006) nachgewiesen werden.

Über die Genese der Selbstkonzepte im Vor- und Grundschulalter ist bekannt und auch aus den oben genannten Erkenntnissen begründet, dass die Kinder noch nicht alle Vergleichsmöglichkeiten (sozial, dimensional, temporal) für ihre Fähigkeitseinschätzungen heranziehen. So fanden Dickhäuser und Galfe (2004), dass Schüler der dritten und vierten Klasse im Gegensatz zu Befunden älterer Schüler und Erwachsener weder Vergleiche zwischen ihren Leistungen in unterschiedlichen Fächern (dimensional) noch Vergleiche zwischen ihren früheren und aktuellen Leistungen (temporal) durchführen. Teilweise bestätigt wurden diese Ergebnisse in einer Übersichtsarbeit zum Internal/External-Frame-of-Reference-Modell (Möller & Köller, 2004). Es zeigte sich hierbei, dass dimensionale Vergleiche frühestens von Viertklässlern durchgeführt werden. Somit scheint das fachspezifische akademische Selbstkonzept zu Beginn der Grundschule zumeist auf sozialen Leistungsvergleichen mit den Mitschülern bzw. auf einer globalen

Einschätzung zu beruhen.

Die Unterschiede der Selbstkonzepteneinschätzung zwischen Jungen und Mädchen waren auch oft Bestandteil von Studien zum Selbstkonzept (Asendorpf & van Aken, 1993; Entwistle et al., 1987; Eccles et al., 1993; Gabriel et al., 2010; Helmke & van Aken, 1995; Jacobs et al., 2002; Marsh et al., 1984; 1991; Valeski & Stipek, 2001; Wigfield & Karpachian, 1991). Bei Betrachtung des globalen Selbstkonzeptes sowie des globalen akademischen Selbstkonzeptes wurden häufig keine Geschlechtsunterschiede gefunden (Jacobs et al., 2002; Marsh et al., 1984; Marsh et al., 1991; 1998). Die Gründe dafür liegen in den verschiedenen Dimensionen, welche das globale Selbstkonzept bzw. das akademische Selbstkonzept bestimmen. Diese sind bei Mädchen und Jungen teilweise unterschiedlich gerichtet sowie unterschiedlich stark ausgeprägt. So zeigten Untersuchungen zur Einschätzung der sportlichen Kompetenz, dass Jungen schon ab dem Kindergartenalter ihre Fähigkeiten signifikant höher beurteilen als Mädchen (Asendorpf & van Aken, 1993; Marsh et al., 1984, 1991; Wigfield et al., 1997). Jacobs et al. (2002) wiesen diesen Geschlechtsunterschied von Beginn der Grundschulzeit überdauernd bis ins junge Erwachsenenalter nach. Dagegen scheinen Mädchen von der ersten bis zur sechsten Klasse ihre musikalischen Fähigkeiten signifikant höher einzuschätzen als Jungen dieser Altersgruppe (Wigfield et al., 1997). Außerdem konnten Marsh et al. (1991) mit Hilfe des facettenreichen Self-Description-Questionnaire zeigen, dass Mädchen zwischen fünf und acht Jahren ein positiveres Selbstkonzept bezüglich ihres Aussehens haben, wobei der Unterschied zwischen den Geschlechtern im Laufe der Entwicklung geringer wird. Im akademischen Bereich wurden im Vorschul- bzw. Grundschulalter hauptsächlich das Selbstkonzept in Mathematik und das verbale Selbstkonzept (Deutsch im

deutschsprachigen Raum, Englisch in englischsprachigen Ländern etc.) betrachtet. Die meisten Studien kommen zu dem Schluss, dass das mathematische Selbstkonzept bei Jungen höher ausgeprägt ist als bei Mädchen (Gabriel et al., 2010; Jacobs et al., 2002; Tiedemann & Faber, 1995; Wagner & Valtin, 2003; Wigfield et al., 1997). In ihrer Metaanalyse konnten Hyde et al. (1990) diesen Effekt schon bei fünfjährigen Kindergartenkindern zeigen. Gleichzeitig betonten sie aber, dass dieser Geschlechtsunterschied in der Altersgruppe zwischen fünf und zehn Jahren mit $d = .08$ noch sehr klein ausfällt und sich erst in der weiteren Entwicklung bis auf $d = .28$ vergrößert. Dies und die eher geringe Varianzaufklärung durch das Geschlecht (Marsh, 1989) sind Gründe dafür, dass in einigen Studien kein Unterschied zwischen Jungen und Mädchen im Kindergarten bzw. zu Beginn der Grundschule gefunden werden konnte (Herbert & Stipek, 2005; Marsh et al., 1991; Valeski & Stipek, 2001). Die Ergebnisse zum verbalen Selbstkonzept sind weniger eindeutig. So fanden Marsh et al. (1991) sowie auch Wigfield et al. (1997) einen signifikanten Vorteil im verbalen Selbstkonzept für die Mädchen bereits ab dem Kindergarten- und frühen Grundschulalter. Der Effekt konnte auch in Untersuchungen ab der 2. Klasse bestätigt werden (Marsh et al., 1984; 1991; Valtin, Wagner & Schwippert, 2005). Dem entgegen stehen Studien, die keine Geschlechtsunterschiede im verbalen Selbstkonzept in der Grundschule nachweisen konnten (Eccles et al., 1993; Nicholls, 1978; Valeski & Stipek, 2001). Jacobs et al. (2002) zeigten in ihrer Längsschnittuntersuchung, dass sich das verbale Selbstkonzept zwischen den Geschlechtern zwar schon während der Grundschulzeit verändert, aber erst zu Beginn der Sekundarstufe berichten Mädchen ein signifikant höheres verbales Selbstkonzept als die Jungen. Zudem gibt es auch eine Studie, in der berichtet wird, dass Jungen in der zweiten Klasse ein höheres verbales Selbstkonzept besitzen als

Mädchen (Wagner & Valtin, 2003). In den weiteren Grundschuljahren ist dieser Unterschied allerdings nicht mehr vorzufinden. Die inhomogenen Resultate könnten zum einen dadurch begründet sein, dass sich das verbale Selbstkonzept aus unterschiedlichen Aspekten zusammensetzt und durch verschiedene Autoren unterschiedlich erfasst wurde. In den meisten Studien zum verbalen Selbstkonzept bei Grundschulern wurde nur nach den Einstellungen zum Lesen gefragt (Eccles et al., 1993; Jacobs et al., 2002; Nicholls, 1978; Valeski & Stipek, 2001; Valtin et al., 2005). Während Marsh et al. (1991) die Fünf- bis Achtjährigen mit dem SDQ-IA nur nach ihren Lesefähigkeiten befragt hat, erweiterte er das verbale Selbstkonzept im SDQ-II durch die Einstellungen zum (muttersprachlichen) Englischunterricht (Marsh, 1989). Wagner und Valtin (2003) dagegen haben das verbale Selbstkonzept auf den Bereich Rechtschreibung limitiert. Ein weiterer Unterschied zwischen den Studien ist in der Anzahl der Items zur Erfassung des akademischen Selbstkonzeptes zu finden. Dies gilt für den mathematischen und auch verbalen Bereich gleichermaßen. So wurden den Kindern in den unterschiedlichen Studien zwischen ein bis zwei (Nicholls, 1978; Valeski & Stipek, 2001; Wagner & Valtin, 2003) und zehn Items (Eccles et al., 1993; Marsh 1989) zur Bearbeitung vorgegeben. Allein die Beachtung der zwei genannten Aspekte schränkt die Vergleichbarkeit der Studien bezüglich eines Geschlechtsunterschieds beim akademischen Selbstkonzept stark ein.

Wie in den oben genannten Studien deutlich wurde, ist die Erfassung des akademischen Selbstkonzeptes bei Kindern im Vor- und Grundschulalter schwieriger und somit auch weniger einheitlich als bei älteren Kindern und Erwachsenen. Die üblicherweise verwendeten paper-pencil-Fragebögen können bei jungen Kindern aufgrund nicht oder nur wenig vorhandener Lesefähigkeiten kaum

eingesetzt werden. Durch die Entwicklung interviewgeleiteter Befragungen ist die Erfassung verschiedener Selbstkonzeptdimensionen auch bei Kindern vom zweiten Schuljahr bis hinunter in den Kindergarten möglich (Entwisle et al., 1987; Helmke, 1991; Marsh et al., 1984; 1991). Sowohl Nicholls (1978) als auch Helmke (1991) wählten in ihren Untersuchungen einen direkteren Weg, um die Einschätzung der Kinder im Vor- und Grundschulalter bezüglich ihrer Leistung zu erfragen: Sie baten die Kinder, sich in eine Reihe von fünf Spielfiguren einzuordnen, die vom besten bis zum schlechtesten Kind als Vergleichsgruppe dienten. So konnte festgestellt werden, wo das Kind seine Position in der Leistungshierarchie sieht. Allerdings sind validierte Instrumente zur Erfassung des Selbstkonzepts von Kindern vor dem Schuleintritt und in den ersten Schuljahren kaum vorhanden. Der von Marsh et al. (1984) entwickelte Self-Description-Questionnaire (SDQ), der bereits bei Fünfjährigen sieben verschiedene Selbstkonzeptdimensionen erfasst, liegt lediglich als englische Fassung vor. Arens et al. (2011) validierten zwar den SDQ-I (Marsh, 1990) in einer deutschen Version, doch die Selbstkonzepterfassung ist hierbei nur in der Altersgruppe zwischen acht und zwölf Jahren möglich. Die Pictorial Scale von Harter und Pike (1984), die eine einfache und somit kindgerechte Erfassung verschiedener Selbstkonzeptskalen ab dem Vorschulalter ermöglichen soll, wurde von Asendorpf und van Aken (1993) in eine deutsche Version übersetzt. Diese Version wurde im Gegensatz zum Original erst ab der zweiten Klasse eingesetzt. Außerdem können keine akademischen Selbstkonzeptdimensionen differenziert erfasst werden, da nur eine Gesamtskala „kognitive Kompetenz“ enthalten ist. Ähnlich verhält es sich mit dem Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen (FEES; Rauer & Schuck, 2004). Dieser enthält zwar eine allgemeine Skala zum

„Selbstkonzept der Schulfähigkeit“, allerdings sind die einzelnen fachspezifischen Selbstkonzepte für Mathematik und Lesen mit nur je einem Item darin repräsentiert. Validierte Erhebungsinstrumente, die Skalen der verschiedenen fachspezifischen Selbstkonzepte beinhalten, sind erst ab Ende der Grundschule zur Durchführung angedacht und normiert (z.B. SESSKO; Schöne et al., 2002; DISK-Gitter von Rost et al., 2007). Zusammenfassend zeigt sich, dass in den meisten Studien mit Grundschulkindern wenige und einfach formulierte Fragen zur Einschätzung akademischer Fähigkeiten Anwendung fanden, denen die Kinder auf mehrstufigen Skalen zustimmen mussten (Dickhäuser & Stiensmeier-Pelster, 2003; Gabriel et al., 2010; Jacobs et al., 2002; Marsh et al., 1991).

Fragestellungen

Die größte Vorhersagekraft für spätere Leistungen innerhalb einer Domäne zeigte sich für das Vorwissen (Chi, 1978; Helmke, 1992), weshalb negative Leistungsentwicklungen schon in frühen Phasen des Bildungswegs beachtet werden sollten. Das Selbstkonzept als wichtiger Faktor in Lern- und Leistungssituationen könnte schon zu Beginn der Schullaufbahn eine mögliche Ursache für Leistungsunterschiede darstellen. Vor allem frühe Leistungsunterschiede zwischen Mädchen und Jungen in den grundlegenden Bereichen Mathematik (z.B. Winkelmann et al., 2008), Lesen und Schreiben (OECD, 2010) könnten auf frühe Selbstkonzeptunterschiede zurückgeführt werden (Self-Enhancement-Ansatz) oder könnten spätere Selbstkonzeptunterschiede hervorrufen (Skill-Development-Ansatz). In beiden Fällen wären Unterschiede zwischen den Geschlechtern ungünstig, da spätestens ab Ende der Grundschule ein reziproker Zusammenhang im Sinne des Reciprocal-Effects-Model (Marsh & Craven, 2006) zwischen den

Leistungen und dem entsprechenden Selbstkonzept angenommen wird (Marsh & Craven, 2006). Mit Hinblick auf die Einfluss- und Fördermöglichkeiten bezüglich des Selbstkonzeptes (Lüdtke, Köller, Marsh & Trautwein, 2005) sollten Geschlechtsunterschiede im akademischen Selbstkonzept so früh wie möglich diagnostiziert werden. Somit ergibt sich eine erste Fragestellung:

1. Sind Unterschiede zwischen Mädchen und Jungen im akademischen Selbstkonzept zu Beginn der Grundschulzeit nachweisbar?

Falls sich Geschlechtsunterschiede im akademischen Selbstkonzept zeigen, könnten diese eventuell aufgrund der Einflussfaktoren zur Genese des Selbstkonzeptes erklärt werden. Wie in den theoretischen Ausführungen erläutert wurde, wirkt zu Beginn der Grundschulzeit hauptsächlich der Skill-Development-Ansatz (Marsh & Craven, 2006), was bedeutet, dass die Leistung das akademische Selbstkonzept beeinflusst und durch Geschlechtsunterschiede in der Leistung somit mögliche Geschlechtsunterschiede im akademischen Selbstkonzept bedingt werden könnten. Zusätzlich zeigte sich, dass die Rückmeldung der Lehrer bedeutender für die Selbstkonzeptgenese ist als die reine Leistung (Dickhäuser & Stiensmeier-Pelster, 2003; Helmke & van Aken, 1995). Die Lehrerrückmeldung wird einerseits durch die erbrachten Leistungen bestimmt, aber auch andere Faktoren, wie z.B. das Geschlecht der Schüler, könnten eine Rolle spielen. Eine zweite Fragestellung lautet also:

2. Welche Variablen haben einen Einfluss auf mögliche Geschlechtsunterschiede im akademischen Selbstkonzept?

4.2 Methode

Stichprobe

An der Untersuchung nahmen 142 Kinder aus neun Klassen vier städtischer Grundschulen in Nordrhein-Westfalen teil. Zum Zeitpunkt der Erhebung waren sie im Durchschnitt 7.09 Jahre alt ($SD = 0.42$ Jahre) und befanden sich am Anfang des zweiten Halbjahres der 1. Klasse. Der Anteil der weiblichen Schüler in der Stichprobe betrug 52.8%.

Material und Vorgehensweise

Zuerst erfolgte die Fähigkeitseinschätzung durch die Lehrer, indem sie allen Schülern Ränge in den Bereichen Rechnen, Lesen und Schreiben zuordneten. Die Anzahl der zu vergebenen Ränge war den Lehrern freigestellt, so dass sich in allen Fähigkeitsbereichen eine Spanne von Rang 1 – bester Schüler – bis Rang 15 – schlechtester Schüler – ergab. Daraufhin wurde das akademische Selbstkonzept mit zwei verschiedenen Methoden erhoben. Dazu wurden die Schüler einzeln durch einen geschulten Testleiter in einen ruhigen Raum innerhalb des Schulgebäudes gebracht. Während der ca. halbstündigen Sitzung beantworteten die Kinder jeweils in den Bereichen Rechnen, Lesen und Schreiben auf einer fünfstufigen Skala (1 – „stimmt gar nicht“ bis 5 – „stimmt genau“) die vier folgenden Items: „Ich bin gut im...“, „...ist leicht für mich“, „Ich bin besser im ... als meine Mitschüler“ und „Ich bin besser im ... als in den anderen Fächern“. Außerdem erfolgte eine Befragung zum Selbstkonzept über eine neu entwickelte Mitschülervergleichsmethode, welche in Studie II detailliert beschrieben wurde. Da der Fokus der Untersuchung im mathematischen Bereich angesiedelt ist, bearbeiteten alle Kinder in einer weiteren halbstündigen Einzelsitzung einen Test zur Erfassung der mathematischen

Fähigkeiten (max. 61 Punkte zu erreichen), der auf dem Modell der mathematischen Entwicklung von Fritz und Ricken (2008) basiert.

4.3 Ergebnisse

Der Selbstkonzeptwert des Mitschülervergleichs ist ein neuer, sich aus den Antworten der Kinder ergebender Rangwert. Die Ermittlung des Rangwerts ist in Studie II beschrieben. Für die Bereiche Rechnen und Lesen ergaben sich Rangwerte zwischen 1 – bester Schüler und 12 – schlechtester Schüler. Beim Schreiben ergaben sich Rangwerte zwischen 1 und 11. Die Rangwerte aus dem Mitschülervergleich sowie die Rangwerte der Lehrereinschätzung wurden aufgrund der besseren Übersichtlichkeit invertiert. Für den anderen Selbstkonzeptwert wurde für jeden Bereich (Rechnen, Lesen, Schreiben) ein Mittelwert aus den vier Selbstkonzeptitems gebildet. Die Reliabilität der Skalen ist mit Cronbach's $\alpha_{\text{Rechnen}} = .74$, $\alpha_{\text{Lesen}} = .82$ und $\alpha_{\text{Schreiben}} = .72$ als zufriedenstellend anzusehen. Die Mittelwerte und Standardabweichungen aller Variablen sind in Tabelle 4.1 dargestellt. Die Datenanalyse bezüglich der Selbstkonzeptwerte sowie des Tests zur Erfassung der mathematischen Fähigkeiten wird mithilfe von t-Tests durchgeführt. Aufgrund der ordinal skalierten Rangwerte der Lehrereinschätzungen und der Selbstkonzeptbestimmung durch den Mitschülervergleich erfolgt die Datenanalyse mit diesen Variablen über nichtparametrische Methoden.

Um Geschlechtsunterschiede im Selbstkonzept zu bestimmen, wurden für das Geschlecht als unabhängige Variable und die Werte des Mitschülervergleichs jeweils in den Bereichen Rechnen, Lesen und Schreiben als abhängige Variablen Mann-Whitney-U-Tests berechnet. Es zeigten sich weder im Rechnen ($U = 2342.50$, $Z = -.421$, $p = .674$) noch im Lesen ($U = 2254.00$, $Z = -.645$, $p = .519$) oder im

Schreiben ($U = 2141.00$, $Z = -1.272$, $p = .203$) signifikante Unterschiede zwischen den Selbstkonzepten von Mädchen und Jungen.

Um Geschlechtsunterschiede im Selbstkonzept mit den herkömmlichen Fragebogenwerten zu bestimmen, wurde für jeden Leistungsbereich ein t-Test mit dem Selbstkonzeptwert als abhängige Variable und Geschlecht als unabhängige Variable durchgeführt. Es zeigt sich weder im Lesen ($t(139) = -.293$, $p = .770$) noch im Schreiben ($t(139) = -.585$, $p = .559$) ein signifikanter Unterschied zwischen den Selbstkonzepten von Mädchen und Jungen. Im Bereich des Rechnens berichten dagegen Mädchen von einem höheren Selbstkonzept ($M = 4.32$, $SD = .608$) als Jungen ($M = 4.04$, $SD = .974$), $t(108.54) = -2.020$, $p = .046$ (t-Test für ungleiche Varianzen).

Eine Erklärung für die Geschlechtsunterschiede im Mathematikselbstkonzept könnten Geschlechtsunterschiede in den Leistungen sein. Allerdings zeigen sich zum einen keine Unterschiede zwischen Mädchen und Jungen im Mathematiktest ($t(138.94) = 1.067$, $p = .288$); zum anderen ergibt die Berechnung des Rangkorrelationskoeffizienten Rho (ρ) nach Spearman (siehe Tabelle 4.2) weder bei Mädchen ($\rho(74) = .033$, $p = .780$) noch bei Jungen ($\rho(67) = .002$, $p = .989$) einen signifikanten Zusammenhang zwischen Mathematiktest und dem zugehörigen Selbstkonzept.

Ein weiterer Einflussfaktor des Selbstkonzepts ist die Lehrereinschätzung. Diese wird u.a. aus der Leistung der Schüler gebildet, was sich in einem signifikanten Zusammenhang zwischen Mathematiktest und Lehrereinschätzung ausdrückt, $\rho(142) = .232$, $p = .005$.

Tabelle 4.1: Mittelwerte und Standardabweichungen für Lehrereinschätzung, Selbstkonzept durch Fragebogenitems, Selbstkonzept durch Mitschülervergleich und Mathematiktest

		Lehrereinschätzung*	SK-Fragebogen	SK- Mitschülervergleich*	Mathematiktest
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Rechnen	Mädchen	11.05 (3.08)	4.32 (.61)	8.31 (2.37)	42.99 (9.80)
	Jungen	10.78 (3.59)	4.04 (.97)	8.41 (2.48)	44.58 (8.0)
	gesamt	10.92 (3.32)	4.19 (.81)	8.36 (2.41)	43.74 (9.0)
Lesen	Mädchen	10.99 (3.27)	4.13 (.96)	8.18 (2.57)	
	Jungen	10.10 (3.64)	4.08 (.92)	7.91 (2.63)	
	gesamt	10.57 (3.47)	4.11 (.94)	8.05 (2.59)	
Schreiben	Mädchen	11.12 (3.29)	4.19 (.79)	7.53 (2.41)	
	Jungen	10.09 (3.25)	4.11 (.85)	7.09 (2.41)	
	gesamt	10.63 (3.30)	4.15 (.82)	7.32 (2.41)	

Anmerkungen. *Werte invertiert, hohe Werte stehen für positive Lehrereinschätzung bzw. hohes Selbstkonzept

Hierbei zeigt sich bei den Mädchen ($\rho(75) = .249, p = .031$) ein vergleichbar hoher Zusammenhang wie bei den Jungen ($\rho(67) = .228, p = .032$, einseitig).

Weiterhin könnten andere Faktoren, wie z.B. das Geschlecht der Schüler die Lehrereinschätzungen beeinflussen. Die Berechnung des Mann-Whitney-U-Tests mit der Lehrereinschätzung als abhängiger und Geschlecht als unabhängiger Variable ergibt jedoch keine Unterschiede in der Leistungseinschätzung durch die Lehrer aufgrund des Geschlechts, $U = 2469.50, Z = -.177, p = .860$. So zeigt die Korrelationsberechnung auch einen vergleichbaren Zusammenhang von Lehrereinschätzung und Selbstkonzept bei Mädchen, $\rho_{Mädchen}(75) = .280, p = .016$ und Jungen, $\rho_{Jungen}(67) = .268, p = .029$.

Tabelle 4.2: Korrelationen (Rangkorrelationskoeffizienten Rho nach Spearman) zwischen Lehrereinschätzung, Leistung im Mathematiktest und dem Selbstkonzept in Mathematik (Fragebogen & Mitschülervergleich) getrennt nach Mädchen und Jungen

	Lehrer- einschätzung	Mathematik- test	SK- Fragebogen	SK- Mitschüler- vergleich
Lehrereinschätzung	----	.228	.268*	.806**
Mathematiktest	.249*	----	.002	.133
SK-Fragebogen	.280*	.033	----	.408**
SK-Mitschülervergleich	.628**	.019	.471**	----

Anmerkungen. * $p < .05$, ** $p < .01$; Werte für Mädchen unten, für Jungen oben; Werte der Lehrereinschätzung & SK-Mitschülervergleich invertiert

4.4 Diskussion

Ziel der vorliegenden Untersuchung war die Feststellung von Geschlechtsunterschieden in verschiedenen Bereichen des akademischen Selbstkonzepts bei Erstklässlern, wobei zwei verschiedene Methoden zur Selbstkonzepterfassung verwendet wurden. Hierbei stellte sich heraus, dass bei der Mitschülervergleichsmethode keine Geschlechtsunterschiede in den Selbstkonzepten hinsichtlich Rechnen, Lesen und Schreiben vorzufinden sind. Das gleiche gilt für die Selbstkonzepte beim Lesen und Schreiben, wenn die Erfassung mit den herkömmlichen Fragebogenmethoden erfolgte. Lediglich im Rechenselbstkonzept wurde ein Unterschied gefunden, wobei Mädchen ihre Fähigkeiten signifikant höher einschätzten als Jungen.

Zusätzlich zur Feststellung von Selbstkonzeptunterschieden zwischen Mädchen und Jungen wurden verschiedene Variablen erfasst, die den gefundenen Unterschied erklären könnten. Allerdings zeigten sich weder im Mathematiktest als Leistungsmaß noch in den Lehrereinschätzungen der Schüler Unterschiede zwischen den Geschlechtern, so dass die Variablen Leistung und Lehrereinschätzung trotz signifikanter Zusammenhänge mit dem Selbstkonzept im Rechnen keine Erklärung für den Geschlechtsunterschied bieten.

Die Leistungseinschätzung der Lehrer z.B. in Form von Noten scheint allerdings nicht von allzu großer Bedeutung für die Genese des akademischen Selbstkonzepts zu sein, wie Dickhäuser und Stiensmeier-Pelster (2003) in ihrer Untersuchung mit Schülern der dritten und vierten Klassenstufe zeigten. Als entscheidender erwies sich die durch die Schüler wahrgenommene Leistungseinschätzung der Lehrer. Diese könnte z.B. durch die Rückmeldung zum Verhalten im Unterricht verursacht werden, was wiederum bei Mädchen und Jungen

unterschiedlich wahrgenommene Leistungseinschätzungen bewirken könnte. Dass die Leistungseinschätzungen z.B. in Form von Noten in einigen Fächern keine Erklärung für Geschlechtsunterschiede im Selbstkonzept liefern, zeigten auch Schilling et al. (2006). Sie erhoben verschiedene akademische Selbstkonzepteingeschätzungen und die entsprechenden Noten von Gymnasiasten der Klassenstufen sieben bis zehn, wobei z.B. die gefundenen Geschlechtsunterschiede im Deutschselbstkonzept durch die Geschlechtsunterschiede in den Noten erklärbar waren, die Geschlechtsunterschiede im Mathematikselbstkonzept allerdings nicht. Die Autoren führten an, dass z.B. die Unterschiede zwischen Mädchen und Jungen in Mathematik auch aufgrund des Stereotypendenkens bedeutsamer anderer Bezugspersonen (z.B. Peers, Eltern) entstehen könnten. Tiedemann (2000) konnte den Einfluss der elterlichen stereotypen Erwartungen und der daraus entstehenden Attributionsmuster für Grundschüler der dritten und vierten Klassen im Fach Mathematik nachweisen. Der in der hier vorliegenden Arbeit gefundene Geschlechtsunterschied kann allerdings nicht durch geschlechtsstereotypische Erwartungen der Eltern erklärt werden, da ja entgegen den Geschlechtsstereotypen Mädchen ihre Fähigkeiten höher einschätzten als Jungen. Interessant wäre eine Untersuchung der Elterneinschätzungen für die Erstklässler, obwohl eine grundlegende Umkehr des Stereotypendenkens der Eltern innerhalb von zwei Schuljahren recht unwahrscheinlich ist. Vielleicht sind die Erwartungen zu Beginn der Schulzeit aber geschlechtsneutral, so dass die Mädchen in der ersten Klasse noch nicht so stark dadurch beeinflusst werden. Zudem könnten auch die Fähigkeitseinschätzungen durch die Mitschüler oder Peers Aufschluss über die Ursachen des gefundenen Unterschieds geben.

Allerdings bleibt zu bedenken, warum sich zwar mit der Fragebogenmethode

ein Unterschied im Rechenselbstkonzept gezeigt hat, nicht aber beim Mitschülervergleich. Bereits in Studie II erwies sich die Methode des Mitschülervergleichs zur Erfassung des Selbstkonzepts bei Erstklässlern als valider im Vergleich zu den herkömmlichen Fragebogenverfahren. Alle Studien, in denen Geschlechtsunterschiede im mathematischen Selbstkonzept gefunden wurden, nutzten Fragebogenverfahren in unterschiedlichen Variationen (Gabriel et al., 2010; Jacobs et al., 2002; Marsh, 1989; Tiedemann & Faber, 1995; Wagner & Valtin, 2003; Wigfield et al., 1997). So kann angenommen werden, dass eine Erklärung von Geschlechtsunterschieden im akademischen Selbstkonzept zu Beginn der Grundschule nicht mehr nötig ist, da dann, wenn die validere Methode zur Erfassung des Selbstkonzepts herangezogen wird, keine Geschlechtsunterschiede gefunden werden, was sich als kongruent mit den erhobenen Leistungsmaßen (Test bzw. Lehrereinschätzung) erweist.

5. Zusammenfassende Diskussion

Im Folgenden werden die Ergebnisse der drei Studien zusammengefasst und im Anschluss daran bezüglich des theoretischen und praktischen Nutzens diskutiert. Abschließend erfolgt ein Ausblick über zukünftige Forschung, die an diese Arbeit anknüpfen könnte.

5.1 Zentrale Ergebnisse

Zunächst einmal sollte die Frage geklärt werden, ob es Unterschiede zwischen Mädchen und Jungen in mathematischen Leistungen der Vor- und Grundschule gibt, da in einigen Untersuchungen Geschlechtsunterschiede erst bei Schülern der weiterführenden Schulen dargestellt werden konnten (Else-Quest et al., 2010; Hyde et al., 1990). Dem entgegen stehen Studien mit jüngeren Kindern, in denen signifikante Unterschiede zwischen Mädchen und Jungen bereits im Grundschulalter nachweisbar waren (Penner & Paret, 2008; van den Heuvel-Panhuizen, 2004; Winkelmann et al., 2008). In Studie I zeigt sich ein signifikanter Vorteil der Jungen zwischen Vorschule und zweiter Klasse. Bei Betrachtung der Mathematikleistungen in den einzelnen Klassenstufen sind die Ergebnisse allerdings inkonsistent. Während im Vorschulalter und der zweiten Klasse keine Geschlechtsunterschiede nachweisbar sind, erreichen Jungen in der ersten Klasse signifikant bessere Werte im Mathematiktest als Mädchen. Der signifikante Unterschied zwischen den Mathematikleistungen von Mädchen und Jungen in der ersten Klasse passt zu einigen bisherigen Befunden (Penner & Paret, 2008; Weinhold Zulauf et al., 2003), wobei Penner und Paret (2008) den Geschlechtsunterschied in dieser Jahrgangsstufe nur für die gesamte Schülerpopulation nachweisen konnten, aber nicht für die Leistungsgruppe der altersgerecht entwickelten Schüler.

Die Betrachtung von verbalen Fähigkeiten in Bezug zu mathematischen Leistungen scheint insofern sinnvoll, da Zusammenhänge dieser beiden Leistungsbereiche immer wieder bestätigt wurden (Aiken, 1971; Krajewski & Schneider, 2009; Möller et al., 2009). Auch in Studie I zeige sich eine signifikante Korrelation zwischen verbalen und mathematischen Leistungen, selbst unter Auspartialisierung der allgemeinen kognitiven Fähigkeiten, $r = .60, p < .001$. Bezüglich der Unterschiede zwischen Mädchen und Jungen ist kein allgemeiner Geschlechtseffekt über die gesamte Altersspanne nachweisbar, da Mädchen und Jungen sowohl in der Vorschulzeit als auch in der ersten Klasse die gleichen Leistungen erbringen. Allerdings steigern Mädchen in der zweiten Klasse ihre verbalen Leistungen weiter, während Jungen sogar unter das Niveau der ersten Klasse zurückfallen. Somit ergibt sich bei den Zweitklässlern ein signifikanter Geschlechtsunterschied. Aufgrund dieser unterschiedlichen Entwicklung der verbalen Leistungen zwischen Mädchen und Jungen kann eine Veränderung der Geschlechtsunterschiede in den mathematischen Leistungen erwartet werden, wenn die verbalen Leistungen berücksichtigt werden. Dies ist insofern der Fall, als dass Mädchen im Vorschulalter unter Einbezug der verbalen Leistungen höhere mathematische Leistungen zeigen als Jungen. Dieser Unterschied ist zwar nicht signifikant, bleibt aber aufgrund der Effektstärke von $d = -.43$ erwähnenswert. Ähnlich verhält es sich mit den mathematischen Leistungen in der ersten Klasse. Allerdings erbringen hier die Jungen bessere Leistungen, was aber ebenso nicht statistisch signifikant ist. Erst in der zweiten Klasse vergrößert sich der Vorteil der Jungen, so dass sie signifikant besser abschneiden als Mädchen. Diese Entwicklung entspricht den Befunden von Weinhold Zulauf et al. (2003), wobei Jungen in dieser Studie bereits ab der ersten Klasse signifikant bessere mathematische Leistungen

erbrachten als Mädchen.

Aufgrund des immer wieder bestätigten Zusammenhangs zwischen akademischen Leistungen und dem entsprechenden Selbstkonzept (Marsh, 1986; 1987; Möller et al., 2009) wurde zur Erklärung der gefundenen Leistungsunterschiede zwischen Mädchen und Jungen das akademische Selbstkonzept betrachtet. Da aber zur Erfassung des akademischen Selbstkonzepts zu Beginn der Grundschulzeit im deutschsprachigen Raum kaum validierte Instrumente vorliegen, wurde in dieser Arbeit zunächst eine neue Methode der Selbstkonzepterfassung entwickelt. Im Gegensatz zu den in Studien üblicherweise verwendeten Fragebogenverfahren (Dickhäuser & Stiensmeier-Pelster, 2003; Gabriel et al., 2010) sollten die Schüler hierbei Fähigkeitsvergleiche mit adaptiv ausgewählten Mitschülern durchführen. Die Ergebnisse der Multitrait-Multimethod-Analyse bestätigen die erfolgreiche Anwendung des neuen Mitschülervergleichs. Für das neue Verfahren zeigt sich im Gegensatz zur Fragebogenmethode eine bessere konvergente Validität aufgrund hoher Korrelationen mit den Fähigkeitseinschätzungen der Lehrerinnen. Da das akademische Selbstkonzept bei Erstklässlern vorrangig auf Rückmeldungen der Lehrkräfte basiert (Helmke & van Aken, 1995), sollten signifikante Korrelationen dieser beiden Variablen aber nachweisbar sein. Für die Bereiche Rechnen und Schreiben ist dies nur beim Mitschülervergleich der Fall. Beim Lesen zeigt sich zwar auch für das Fragebogenverfahren eine signifikante Korrelation mit der Lehrereinschätzung, allerdings ist diese eher schwach und auch deutlich geringer als die Korrelation beim Mitschülervergleich. Die Ergebnisse für den Bereich Lesen könnten dadurch bedingt sein, dass die Kinder auch im außerschulischen Kontext mehr Rückmeldungen über ihre Lesefähigkeit erhalten als Rückmeldungen im Rechnen oder Schreiben,

wodurch sie erfahrener darin sind, ihre Lesefähigkeiten auch im Vergleich zu den anderen Kindern einzuschätzen. Für beide Verfahren zur Selbstkonzepterfassung gilt, dass sich die drei Selbstkonzeptbereiche Rechnen, Lesen und Schreiben nicht eindeutig voneinander trennen lassen, was für eine geringe diskriminante Validität spricht. Dies war aber aufgrund der bisherigen Selbstkonzeptforschung mit jungen Kindern zu erwarten, da sich die einzelnen Selbstkonzeptdimensionen erst im Laufe der Entwicklung besser differenzieren lassen (Marsh et al., 1984; 1991; Shavelson et al., 1976). Aufgrund der adaptiven Auswahl der Vergleichsmitschüler basierend auf der Lehrereinschätzung könnten die hohen Korrelationen allein dadurch zustande gekommen sein. Dies kann allerdings durch die Analyse mit Zufallswerten als unwahrscheinlich herausgestellt werden, da die Berechnungen von artifiziellen Mitschülervergleichswerten mit den Lehrereinschätzungen oft Nullkorrelationen ergeben bzw. die Korrelationen in den meisten Fällen niedriger ausfielen als bei den Berechnungen mit den Originaldaten. Zusammenfassend kann gesagt werden, dass durch die Mitschülervergleichsmethode das akademische Selbstkonzept von Erstklässlern präziser abgebildet wird als durch das üblicherweise angewendete Fragebogenitem des sozialen Vergleichs.

In der dritten Studie wurden erneut beide Methoden zur Selbstkonzepterfassung bei Erstklässlern eingesetzt, um mögliche Geschlechtsunterschiede im akademischen Selbstkonzept in den Bereichen Rechnen, Lesen und Schreiben nachzuweisen. Beim Mitschülervergleich kann kein Unterschied zwischen Mädchen und Jungen in den akademischen Selbstkonzepten des Rechnens, Lesens und Schreibens festgestellt werden. Dies gilt ebenso für die Bereiche Lesen und Schreiben, wenn das akademische Selbstkonzept mit den üblicherweise angewandten Fragebogenmethoden erfasst wird. Im Bereich des

Rechnens schätzen Mädchen dagegen ihre Fähigkeiten signifikant höher ein als Jungen. Um diesen Unterschied möglicherweise erklären zu können, wurden weitere Variablen, ein Mathematiktest als Leistungsmaß und Fähigkeitseinschätzungen durch die Lehrkräfte, erhoben. Es zeigen sich zwar signifikante Korrelationen zwischen dem Rechenselbstkonzept und der Lehrereinschätzung, allerdings können keine Unterschiede zwischen Mädchen und Jungen in dieser Variablen sowie auch beim Mathematiktest nachgewiesen werden, sodass eine Erklärung des Geschlechtsunterschieds im Rechenselbstkonzept nicht dadurch begründet werden kann.

5.2 Theoretischer Nutzen

Aus den Ergebnissen der einzelnen Studien ergeben sich einige Erträge, die für den wissenschaftlichen Bereich als nützlich anzusehen sind. So zeigt sich, dass die Befunde aus Studie I die Ergebnisse der Untersuchung von Winkelmann et al. (2008) sowie IGLU (Walther et al., 2003) ergänzen. In den beiden letztgenannten Studien wurden signifikante Vorteile für Jungen der dritten und vierten Klasse in mathematischen Fähigkeiten gefunden. Die Ergebnisse aus Studie I zeigen, dass Geschlechtsunterschiede nicht erst zum Ende der Grundschulzeit nachweisbar sind, sondern bereits in der zweiten Klasse. Somit kann für Kinder in Deutschland eine Entwicklung der Geschlechtsunterschiede im mathematischen Bereich von der Vorschulzeit bis zum Ende der Grundschule dargestellt werden. Zusätzlich stützen die Befunde der Studie I die Ergebnisse von Weinhold Zulauf et al. (2003). In beiden Studien zeigte sich ein leichter, aber statistisch nicht signifikanter Vorteil der Mädchen im Kindergarten, der sich kurze Zeit später insofern verändert, als dass Jungen signifikant bessere mathematische Leistungen erbringen. Unterschiedlich ist

allerdings der Zeitpunkt, an dem die Geschlechtsunterschiede erstmals auftreten. Während Jungen in der Schweiz und den USA bereits ab der ersten Klasse bessere Testleistungen aufweisen als Mädchen (Penner & Paret, 2008; Weinhold Zulauf et al., 2003), zeigt sich dieser Vorteil bei den deutschen Schülern dieser Studie erst ab der zweiten Klasse. Der Unterschied könnte darin begründet sein, dass in Studie I nur mathematisch durchschnittliche Schüler betrachtet wurden, während in den anderen Studien mit einer repräsentativen Stichprobe der gesamten Schülerpopulation gearbeitet wurde. Neben der ausgewählten Leistungsgruppe wurden in Studie I die verbalen Fähigkeiten als eine weitere Variable in Bezug auf die mathematischen Fähigkeiten betrachtet. Die Ergebnisse bezüglich der Geschlechtsunterschiede verändern sich, je nachdem, ob die verbalen Fähigkeiten in die Berechnung einbezogen werden oder nicht. So sind also auch Unterschiede zu Studien zu erwarten gewesen, in denen die verbalen Fähigkeiten nicht berücksichtigt wurden (Penner & Paret, 2008; Weinhold Zulauf et al., 2003). Weiterhin wurden die mathematischen Fähigkeiten in Studie I mit einem Test erhoben, der auf einem neuen theoretischen Modell zur mathematischen Entwicklung basiert (Fritz & Ricken, 2008). Da dieses Modell vom Erwerb aufeinander aufbauender Konzepte in der mathematischen Entwicklung ausgeht, können Unterschiede in den Ergebnissen zwischen Studie I und Studien, die z.B. den curriculumsbasierten Ansatz zur Erfassung mathematischer Fähigkeiten gewählt haben, begründet sein (Tiedemann & Faber, 1994). Denkbar wäre zudem, dass der Mathematiktest in Studie I die verbalen Fähigkeiten mehr oder anders beansprucht als bei der Erfassung mathematischer Fähigkeiten in anderen Studien, sodass hierbei die verbalen Fähigkeiten besonderer Berücksichtigung benötigen als bei anderen Tests. In jedem Fall ist Studie I aufgrund der in Deutschland bzw. im deutschsprachigen Raum nur

wenig vorhandenen Studien zu Geschlechtsunterschieden in Mathematik im Vor- und Grundschulalter als Bereicherung dieses Forschungsgebiets anzusehen. Als gewinnbringend ist vor allem anzumerken, dass die in Studie I berücksichtigten Einflussfaktoren (verbale Fähigkeiten, Art der Erfassung mathematischer Fähigkeiten, spezielle Leistungsgruppe) Erklärungen für die inkonsistenten Ergebnisse in Studien zu Geschlechtsunterschieden im mathematischen Bereich bieten.

Ebenso inkonsistent sind die Befunde zu Geschlechtsunterschieden im akademischen Selbstkonzept bei Kindern im Vor- und Grundschulalter. Wahrscheinlich auch aufgrund der eingeschränkten Möglichkeit, das akademische Selbstkonzept bei jungen Kindern zu erfassen (z.B. aufgrund geringer Lesefähigkeiten), fehlen standardisierte deutschsprachige Verfahren für den Beginn der Grundschule. In vielen Studien zum akademischen Selbstkonzept in diesem Altersbereich wurden eigens formulierte Fragen verwendet, die aus Fragebögen für ältere Kinder oder Erwachsene vereinfacht übernommen wurden. Hierbei unterscheiden sich die Studien aber einerseits in der Anzahl der gestellten Fragen zum Selbstkonzept und andererseits in den angesprochenen Bezugsnormen (soziale, individuelle oder dimensionale Vergleiche, globale Aussagen). Das in Studie II vorgestellte neue Verfahren zur Erfassung des akademischen Selbstkonzepts bei Kindern zu Beginn der Grundschulzeit berücksichtigt dagegen die noch andauernde kognitive Entwicklung. Die Kinder müssen sich z.B. nicht, wie es bei Helmke (1991) oder den typischen Fragebogenverfahren (Gabriel et al., 2010; Marsh et al., 1991) der Fall ist, mit einer großen Gruppe vergleichen, was in diesem Altersbereich sehr fehleranfällig ist (Nicholls, 1978), sondern erhalten einzeln konkrete Mitschüler für einen Fähigkeitsvergleich genannt. Das Verfahren von Nicholls (1978) und auch

Helmke (1991) bietet allerdings eine Erleichterung für die Kinder, dadurch dass sie keine abstrakte Einschätzung auf einer Skala vornehmen müssen, sondern ihre Fähigkeitseinschätzung anhand eines konkreten und anschaulichen Aufbaus vornehmen können. Die Idee der Rangzuordnungen wurde deshalb für den Mitschülervergleich übernommen. Zudem wird auch nur der soziale Vergleich berücksichtigt, da die dimensional und temporalen Vergleiche erst zum Ende der Grundschulzeit durchgeführt werden (Dickhäuser & Galfe, 2004; Helmke, 1991). Somit bietet der neue Mitschülervergleich durch die erweiterte Berücksichtigung der kognitiven Entwicklung und der Genese des Selbstkonzepts im Grundschulalter theoretische Vorteile gegenüber den herkömmlich angewendeten Fragebogenmethoden. Zudem zeigen die Befunde aus Studie II, dass die theoretischen Vorteile scheinbar auch zu valideren Ergebnissen der Selbstkonzepterfassung führen. Somit sollte der Mitschülervergleich eher Anwendung finden als Fragebogenmethoden.

Der Vergleich der zwei Methoden zur Erfassung des akademischen Selbstkonzepts wurde in Studie III weiter betrachtet, um Erkenntnisse über Geschlechtsunterschiede beim akademischen Selbstkonzept sehr junger Schüler zu erhalten, die vor allem für diesen Altersbereich nur wenig vorhanden sind. Lediglich im Rechen selbstkonzept konnten signifikante Unterschiede zwischen Mädchen und Jungen gefunden werden, wenn das Selbstkonzept mit der Fragebogenmethode erfasst wurde. Zusätzlich erhobene Variablen (Leistung, Lehrereinschätzung) konnten keine Erklärung für die höhere Fähigkeitseinschätzung der Mädchen bieten. Eine Erklärung des Geschlechtsunterschieds ist eventuell auch gar nicht nötig, da die Erfassung des akademischen Selbstkonzepts mit der valideren Methode des Mitschülervergleichs als beachtenswerter anzusehen ist. Da dort keine Unterschiede

zwischen Mädchen und Jungen nachgewiesen werden konnten, kann davon ausgegangen werden, dass Mädchen wie Jungen zu Beginn der Grundschulzeit ihre akademischen Fähigkeiten gleich einschätzen. Die bisher gefundenen inkonsistenten Ergebnisse bezüglich der Geschlechtsunterschiede beim akademischen Selbstkonzept junger Kinder könnten also auf die unterschiedlichen Erhebungsmethoden zurückgeführt werden. In den Studien von Valeski und Stipek (2001) sowie Herbert und Stipek (2005) wurde zum Beispiel auf Fragebogenitems des sozialen Vergleichs verzichtet, so dass im Gegensatz zu anderen Studien, in denen die Kinder einen sozialen Vergleich durchführen sollten (Tiedemann & Faber, 1995; Wagner & Valtin, 2003), keine Geschlechtsunterschiede zu Beginn der Grundschule nachgewiesen werden konnten.

5.3 Praktischer Nutzen

Neben dem empirischen Erkenntnisgewinn für den Forschungsbereich der mathematischen Entwicklung und zugehöriger Selbstkonzepte im Vor- und Grundschulalter, können auch praktische Implikationen aus den Studien I bis III abgeleitet werden. Wie bereits im Theorieteil erwähnt wurde, ist die Berufswahl stark geschlechtsstereotypisch geprägt. Dies kann u.a. mit dem reziproken Zusammenhang zwischen erbrachten Leistungen und den entsprechenden Selbstkonzepten erklärt werden. Im Sinne des Skill-Development-Ansatzes wirken die Leistungen auf das Selbstkonzept. Sind die Leistungen zwischen Mädchen und Jungen unterschiedlich, so entwickeln Mädchen und Jungen auch unterschiedliche Selbstkonzepte in den entsprechenden Bereichen. Der im Self-Enhancement-Ansatz postulierte Einfluss des akademischen Selbstkonzepts auf die Leistungen verstärkt die Unterschiede während der Schullaufbahn. Da auch aufgrund des

Fachkräftemangels in technischen Berufen eine geschlechtsneutrale Berufswahl angestrebt werden sollte, bietet die Kenntnis von Geschlechtsunterschieden Ansatzpunkte zur Veränderung. In Studie I wurde deutlich, dass Geschlechtsunterschiede bereits in der ersten bzw. zweiten Klasse (unter Berücksichtigung der verbalen Fähigkeiten) vorhanden sind. Selbst die nicht signifikanten Ergebnisse in der Vorschule und der zweiten Klasse (ohne Berücksichtigung der verbalen Fähigkeiten) können als kleine Effekte angesehen werden, die sogar höher ausfallen ($d = .20/.28$) als der signifikante Geschlechtsunterschied in der Mathematikleistung von Viertklässlern in der TIMS-Studie 2007 ($d = .12$). Studie III zeigte, dass die Einschätzung der eigenen akademischen Fähigkeiten für Mädchen und Jungen der ersten Klasse noch gleich hoch ist. Trotzdem sollte v.a. das Wissen über frühe Leistungsunterschiede zwischen Mädchen und Jungen einerseits auf bildungspolitischer Ebene und andererseits auch unterrichtsbegleitend durch Lehrer genutzt werden (Lüdtke et al., 2005), um der Manifestation geschlechtsstereotyper Leistungs- und Selbstkonzeptunterschiede im Laufe der Schulzeit entgegenzuwirken. In ihrer Metaanalyse bezüglich Interventionen zur Verbesserung akademischer Selbstkonzepte bieten O'Mara, Marsh, Craven und Debus (2006) verschiedene Ansatzpunkte, damit Lehrer dazu beitragen können, dass sich das akademische Selbstkonzept und daraufhin die Leistungen der Schüler positiv entwickelt. So zeigte sich, dass Interventionen zur Prävention negativer akademischer Selbstkonzepte fast ebenso wirksam sind wie Interventionen bei bereits vorhandenen Problemen. Dies ist für Lehrer von Schülern zu Beginn der Schulzeit insofern wichtig, da sich zwar z.B. noch keine negativen mathematischen Selbstkonzepte bei Mädchen im Gegensatz zu Jungen zeigen, aber einem späteren Auftreten des Geschlechtsunterschieds bereits entgegengewirkt

werden kann. Eine zusätzliche Erkenntnis der Metaanalyse ist außerdem, dass Lehrer am geeignetsten sind, um eine Verbesserung des akademischen Selbstkonzepts zu bewirken. Zudem können Lehrer die Fähigkeiten bzw. Leistungen ihrer Schüler adäquat einschätzen (Tiedemann & Faber, 1994), was die hohen Korrelationen zwischen Lehrereinschätzung und Testwert für Mathematik in Studie II und III zusätzlich bestätigten. Mit der in Studie II vorgestellten neuen Methode des Mitschülervergleichs erhalten Lehrer nun auch für den Anfangsunterricht eine valide Möglichkeit, das akademische Selbstkonzept ihrer Schüler zu erfassen. Somit können sie überprüfen, ob schlechte Leistungen bzw. Leistungsunterschiede zwischen den Geschlechtern eventuell auf negative akademische Selbstkonzepte zurückzuführen sind. Falls dies der Fall ist, haben sie im Sinne des Self-Enhancement-Ansatzes die Möglichkeit Leistungsverbesserungen durch Selbstkonzepterhöhung zu bewirken.

Dabei sollte darauf geachtet werden, dass nicht nur der vordergründig zu verbessernde Bereich bearbeitet wird, sondern auch Einflussfaktoren beachtet werden. Wie in Studie I erkennbar ist, verändern sich die Unterschiede von Mädchen und Jungen im mathematischen Bereich, wenn verbale Fähigkeiten mit in die Analyse einbezogen werden. Auch wenn Lehrer keine Unterschiede zwischen mathematischen Leistungen von Mädchen und Jungen feststellen, sollten sie z.B. trotzdem die verbalen Fähigkeiten berücksichtigen, um auszuschließen, dass die Mädchen geringere mathematische Fähigkeiten lediglich mit besseren verbalen Fähigkeiten ausgleichen.

Wie in den Ausführungen deutlich geworden ist, findet der neue Mitschülervergleich vor allem innerhalb einzelner Klassen Anwendung, da die neu entstandenen einzelnen Rangwerte auch nur innerhalb der zugehörigen Klasse

interpretiert werden können. Allerdings zeigen Studie II und III, dass der Mitschülervergleich durchaus auch bei empirischen Forschungsfragen angewendet werden kann.

5.4 Zusammenfassung und Ausblick

Aufgrund der Ergebnisse in Studie I muss der Aussage von Else-Quest et al. (2010) oder auch Hyde et al. (1990), dass Geschlechtsunterschiede in Mathematik so gut wie nicht mehr vorhanden wären, widersprochen werden. Als globale Aussage entsprechend einem internationalen Durchschnitt ist dies vielleicht zutreffend, aber zumindest bei deutschen Schülern sind Unterschiede zwischen Mädchen und Jungen in Mathematik frühzeitig in der Schullaufbahn nachweisbar und sollten deshalb Beachtung finden. Auch in internationalen Vergleichsstudien wie PISA (z.B. OECD, 2009) zeigen sich große Unterschiede zwischen einzelnen Staaten bezüglich mathematischer Kompetenzen von Mädchen und Jungen. Guiso et al. (2008) fanden mit der *gender stratification hypothesis* eine Erklärung für diese internationalen Differenzen. Dabei gehen sie davon aus, dass die Geschlechtsunterschiede in Mathematik vom Status der Frau innerhalb der entsprechenden Kultur bzw. innerhalb des Staates abhängig sind. Somit sollte der Fokus von Studien zu Geschlechtsunterschieden in Mathematik eher auf dem nationalen Zustand liegen als internationale Durchschnittswerte zu bilden, die für die einzelnen Staaten von geringer Bedeutung sind. Zudem scheinen verschiedene Variablen, wie z.B. verbale Fähigkeiten, die Unterschiede zwischen Mädchen und Jungen in Mathematik zu beeinflussen. In zukünftigen Studien sollten die verbalen Fähigkeiten auch im mathematischen Forschungsbereich stärkere Berücksichtigung finden, zumal die im Abschnitt „Theoretischer Nutzen“ angesprochene mögliche Interaktion zwischen

verbalen Fähigkeiten und Art der Erfassung mathematischer Leistungen bzw. Fähigkeiten überprüfenswert wäre. Für detailliertere Aussagen zu Geschlechtsunterschieden in Mathematik ist außerdem die Beachtung unterschiedlicher Leistungsgruppen von Vorteil. Beim Vergleich von Studie I und III zeigt sich nämlich, dass je nach Auswahl der Stichprobe unterschiedliche Ergebnisse auftreten. In Studie I konnte in der Stichprobe aus mathematisch altersadäquat entwickelten Erstklässlern ein signifikanter Vorteil der Jungen nachgewiesen werden, während in Studie III alle Erstklässler der teilnehmenden Klassen Teil der Stichprobe waren und kein Geschlechtsunterschied in Mathematik festgestellt wurde. Unter Verwendung einer größeren Stichprobe für Studie III und der Einteilung der Schüler in Leistungsgruppen sind vergleichbare Ergebnisse bezüglich mathematischer Geschlechtsunterschiede in beiden Studien anzunehmen.

Auch in Studie II wäre eine größere und vor allem differenziertere Stichprobe zur Absicherung der Befunde sinnvoll. Eine Erhebung in unterschiedlichen Regionen, Bundesländern oder sogar Staaten würde die allgemeine Gültigkeit des Mitschülervergleichs bestätigen. Um zu überprüfen, bis zu welchem Altersbereich die Erfassung des akademischen Selbstkonzepts durch den Mitschülervergleich valider ausfällt als durch Fragebogenmethoden, wären längsschnittliche Untersuchungen über die gesamte Grundschulzeit von Vorteil. Da die Vor- und Nachbereitung für die Durchführung des Mitschülervergleichs recht zeitaufwendig ist, könnte ein Computerprogramm bei der Auswahl der Vergleichsmitschüler und der Berechnung der Selbstkonzeptwerte behilflich sein. Dies könnte nicht nur die Arbeit bei weiteren empirischen Untersuchungen erleichtern, sondern auch die Anwendung des Verfahrens für Lehrer vereinfachen.

In allen Studien der vorliegenden Arbeit wurde ein querschnittliches

Untersuchungsdesign zur Beantwortung der Fragestellungen verwendet. Um aber vor allem die Entwicklung von Geschlechtsunterschieden in Mathematik bzw. dem akademischen Selbstkonzept vom Vorschulalter bis zum Ende der Grundschulzeit umfassend abbilden zu können, sollten eher Längsschnittdaten erhoben werden. Eine Längsschnittstudie, in der sowohl die mathematische als auch die Selbstkonzeptentwicklung betrachtet wird, könnte zusätzliche Erkenntnisse über die Zusammenhänge zwischen Leistung und dem entsprechenden Selbstkonzept im Sinne des Reciprocal-Effects-Model (Marsh & Craven, 2006) für den sehr jungen Altersbereich bringen. Dabei sollten zusätzliche Variablen erhoben werden, die mögliche Geschlechtsunterschiede im akademischen Selbstkonzept erklären können. Die in Studie III erfasste Fähigkeitseinschätzung durch den Lehrer sowie auch die Testleistung konnten keine Erklärung für den Geschlechtsunterschied im Rechenselbstkonzept bieten. Auch Schilling et al. (2006) fanden heraus, dass die Lehrereinschätzung z.B. in Form von Noten zwar die Geschlechtsunterschiede im Deutschselbstkonzept, aber nicht im Mathematikselbstkonzept bei Schülern der Sekundarstufe begründeten. Deshalb könnten geschlechtsstereotype Einstellungen bedeutsamer Bezugspersonen (Freunde, Eltern) oder auch die von Schülern wahrgenommenen Lehrereinschätzungen vielversprechende Ansatzpunkte darstellen, um Geschlechtsunterschiede im akademischen Selbstkonzept und eventuell entsprechender Unterschiede in den Leistungen der Schüler erklären zu können (Dickhäuser & Stiensmeier-Pelster, 2003; Tiedemann, 2000). Spätestens nach Absicherung der Existenz von Leistungs- bzw. Selbstkonzeptunterschieden und deren Ursachen in empirischen Arbeiten sollten Lehrer und Eltern auf diese Fakten aktiv in Aus- und Weiterbildung sowie Informationsveranstaltungen aufmerksam gemacht werden. Gleichzeitig sollten sie Möglichkeiten an die Hand

bekommen, wie sie die Leistungs- und Selbstkonzeptentwicklung positiv beeinflussen können. Interventionen zur Verbesserung von Mathematikfähigkeiten (z.B. Fritz & Gerlach, 2011) oder zur Erhöhung des akademischen Selbstkonzepts (O'Mara et al., 2006) sind bereits vorhanden. Allerdings könnte ein kombiniertes Programm speziell für Kinder im Vor- und Grundschulalter empirisch begleitet entwickelt werden.

6. Literatur

- Aiken, L.R. (1971). Verbal factors in mathematics learning: A review of research. *Journal for Research in Mathematics Education*, 2, 304-313.
- Ambady, N., Shih, M., Kim, A. & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 12, 385-390.
- Arens, A.K., Trautwein, U. & Hasselhorn, M. (2011). Erfassung des Selbstkonzepts im mittleren Kindesalter: Validierung einer deutschen Version des SDQ I. *Zeitschrift für Pädagogische Psychologie*, 25, 131-144.
- Asendorpf, J.B. & van Aken, M.A.G. (1993). Deutsche Versionen der Selbstkonzeptskalen von Harter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 64- 86.
- Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C. E. Granrud (Hrsg.), *Visual perception and cognition in infancy* (S. 265-316). Hillsdale, N.J.: Erlbaum.
- Baillargeon, R. & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development*, 62, 1227-1246.
- Baron-Cohen, S. (2003). *The essential difference: The truth about the male and female brain*. New York: Basic Books.
- Barth, H., Kanwisher, N. & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, 86, 201–221.
- Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N. & Spelke, E. (2006).

- Non-symbolic arithmetic in adults and young children. *Cognition*, 98, 199–222.
- Beilock, S. L., Gunderson, E. A., Ramirez, G. & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences, USA*, 107, 1060-1063.
- Berg, D. H. (2008). Working memory and arithmetic calculation in children: The contributory roles of processing speed, shortterm memory, and reading. *Journal of Experimental Child Psychology*, 99, 288–308.
- Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C. & Walther, G. (Hrsg.) (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G., & Valtin, R. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Brannon, E. M. & Van de Walle, G. A. (2001). The development of ordinal numerical competence in young children. *Cognitive Psychology*, 43, 53–81.
- Burba, D. (2006). *Leistungen von Mädchen und Jungen bei PISA 2003—bedeutsame Unterschiede?* Unveröffentlichte Dissertation, Christian-Albrechts-Universität zu Kiel.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

- Calsyn, R. & Kenny, D. (1977). Self-concept of ability and perceived evaluations by others: Cause or effect of academic achievement. *Journal of Educational Psychology*, 69, 136–145.
- Cattel, R.B., Weiß, R.H. & Osterland, J. (1997). *CFT 1. Grundintelligenztest Skala 1* (5., revidierte Aufl.). Göttingen: Hogrefe.
- Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Hrsg.), *Children's thinking: What develops?* (S. 73–96). Hillsdale: Erlbaum.
- Cohen J. (1969): *Statistical power analysis for the behavioral sciences*. Academic Press, New York, London.
- Cooley, C. H. (1902). *Human nature and the social order*. New York: Charles Scribner's Sons.
- Davis, H. & Carr, M. (2002). Gender differences in mathematics strategy use: the influence of temperament. *Learning and Individual Differences*, 13, 83–95.
- Dickhäuser, O. (2006). Fähigkeitsselbstkonzepte: Entstehung, Auswirkung, Förderung. *Zeitschrift für Pädagogische Psychologie*, 20, 5-8.
- Dickhäuser, O. & Galfe, E. (2004). Besser als..., schlechter als... - Leistungsbezogene Vergleichsprozesse in der Grundschule. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 36, 1-9.
- Dickhäuser, O. & Stiensmeier-Pelster, J. (2003). Wahrgenommene Lehrereinschätzungen und das Fähigkeitsselbstkonzept von Jungen und Mädchen in der Grundschule. *Psychologie in Erziehung und Unterricht*, 50,

182-190.

Eccles, J. S., Wigfield, A., Harold, R. & Blumenfeld, P. B. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 64, 830–847.

Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103–127.

Entwisle, D. R., Alexander, K. L., Pallas, A. M. & Cadigan, D (1987). The emergent academic self-image of first graders: Its response to social structure. *Child Development*, 58, 1190-1206.

Feigenson, L. & Carey, S. (2003) Tracking individuals via object-files: evidence from infants' manual search. *Developmental Science*, 6, 568–584.

Feigenson L., Carey S. & Hauser M. (2002). The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychological Science*. 13, 150–156.

Feigenson, L., Dehaene, S. & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307-314.

Fend, H. (1991). *Identitätsentwicklung in der Adoleszenz. Lebensentwürfe, Selbstfindung und Weltaneignung in beruflichen, familiären und politisch-weltanschaulichen Bereichen. (Bd. II)*. Bern: Huber.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.

- Fritz, A. & Ricken, G. (2008). *Rechenschwäche*. München: UTB: Reinhardt.
- Gabriel, K., Kastens, C., Poloczek, S., Schoreit, E. & Lipowsky, F. (2010).
Entwicklung des mathematischen Selbstkonzepts im Anfangsunterricht - Der
Einfluss des Klassenkontextes. *Zeitschrift für Grundschulforschung*, 3, 65–
82.
- Gallagher, A. M., Levin, J. Y. & Cahalan, C. (2002). *Cognitive patterns of gender
differences on mathematics admissions tests*. (ETS Research Report No. 02-
19). Princeton, NJ: Educational Testing Service.
- Geary, D. C. (1998). *Male, female: The evolution of human sex differences*.
Washington, DC: American Psychological Association.
- Gerlach, M. & Fritz, A. (2011). *Die Biene Mina und der Maulwurf. Ein
Rechenttraining für den Kindergarten*. Berlin: Cornelsen.
- Grimm, H. & Schöler, H. (1978). *Der Heidelberger Sprachentwicklungstest H-S-E-
T*. Göttingen: Hogrefe. (2. korrigierte Aufl. 1991).
- Guiso, L., Monte, F., Sapienza, P. & Zingales, L. (2008). Culture, gender, and math.
Science, 320, 1164–1165.
- Halpern, D. (2000). *Sex differences in cognitive abilities* (3rd edition.). Mahwah, NJ:
Erlbaum.
- Halpern, D. F., Benbow, C., Geary, D., Gur, D., Hyde, J. & Gernsbacher, M.A.,
(2007). The science of sex-differences in science and mathematics.
Psychological Science in the Public Interest, 8, 1-51.
- Hampson, E. (1990). Variations in sex-related cognitive abilities across the

- menstrual cycle. *Brain & Cognition*, 14, 26-43.
- Hannover, B. (1991). Zur Unterrepräsentanz von Mädchen in Naturwissenschaften und Technik: Psychologische Prädiktoren der Fach- und Berufswahl. *Zeitschrift für Pädagogische Psychologie*, 5, 169-186.
- Harter, S. & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development*, 55, 1969-1982.
- Heine, C., Kerst, C. & Sommer, D. (2007) *Studienanfänger im Wintersemester 2005/06*. Hannover: HIS, Hochschul-Informations-System-GmbH.
- Helmke, A. (1991). Entwicklung des Fähigkeitsselbstbildes vom Kindergarten bis zur dritten Klasse. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung. Ein Resumee der Längsschnittforschung* (S. 83-99). Stuttgart: Enke.
- Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen*. Göttingen: Hogrefe.
- Helmke, A. (1998). Vom Optimisten zum Realisten? Zur Entwicklung des Fähigkeitsselbstbildes vom Kindergarten bis zur 6. Klassenstufe. In F.E. Weinert (Hrsg.), *Entwicklung im Kindesalter* (S. 115-132). Weinheim: Beltz.
- Helmke, A., Hosenfeld, I., Groß Ophoff, J., Halt, A.C., Hochweber, J., Isaac, K., Koch, U. & Scherthan, F. (2006). *Ergebnisbericht VERA 2005 NRW*. Universität Koblenz - Landau, Campus Landau.
- Helmke, A., Hosenfeld, I., Groß Ophoff, J., Hochweber, J., Isaac, K., Koch, U., Scherthan, F. & Wagner, S. (2005). *Ergebnisbericht VERA 2004 NRW*. Universität Koblenz - Landau, Campus Landau.

- Helmke, A. & van Aken, M.A.G. (1995). The causal ordering of academic achievement and selfconcept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, 87, 624-637.
- Herbert, J. & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Applied Developmental Psychology*, 26, 276–295.
- Hespos, S. J. & Baillargeon, R. (2001). Knowledge about containment events in very young infants. *Cognition*, 78, 207-245.
- Hosenfeld, I., Köller, O. & Baumert, J. (1999). Why sex differences in mathematics achievement disappear in German secondary schools: A reanalysis of the German TIMSS-data. *Studies in Educational Evaluation*, 25, 143–161.
- Hunting, R.-P. (2003). Part-whole number knowledge in preschool children. *Journal of Mathematical Behavior*, 22, 217–235.
- Huston, A. (1987). The development of sex typing: Themes from recent research. In S. Chess & A. Thomas (Hrsg.), *Annual progress in child psychology and child development* (S. 169–186). New York: Brunner/Mazel.
- Hyde, S. H., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139–155.
- Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect, a meta-analysis. *Psychology of Women Quarterly*, 14, 299-324.
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. & Williams, C. (2008). Gender

similarities characterize math performance. *Science*, 321, 494–495.

Institut der deutschen Wirtschaft Köln (2008). *Wachstums- und Fiskaleffekte von Maßnahmen gegen Fachkräftemangel in Deutschland - Bildungsökonomische Analyse und politische Handlungsempfehlungen insbesondere im MINT-Bereich*. Köln: Bundesministerium für Wirtschaft und Technologie. [Online] Verfügbar unter:
<http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/fiskaleffekte-fachkraeftemangel,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf> [27.08.2012].

Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S. & Wigfield, E. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73, 509-527.

James, W. (1890). *The principles of psychology*. New York: Holt, Rinehart & Winston.

Jansen, H., Mannhaupt, G., Marx, H. & Skowronek, H. (2002). *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC)*. Göttingen: Hogrefe.

Kammermeyer, G. & Martschinke, S. (2006). Selbstkonzept, Lernfreude und Leistungsangst und ihr Zusammenspiel im Anfangsunterricht. In A. Schröder-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang* (S. 125-139). Wiesbaden: VS Verlag für Sozialwissenschaften.

Kellman, P. J. (1993). Kinematic foundations of infant visual perception. In C. E.

- Granrud (Hrsg.), *Visual perception and cognition in infancy. Carnegie-Mellon Symposia on Cognition* (S. 121-173). Hillsdale, NJ; Erlbaum.
- Kenney-Benson, G., Pomerantz, E., Ryan, A. & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, 42, 11–26.
- Kimura D. (2002). Sex hormones influence human cognitive pattern. *Neuroendocrinology Letters, Special issue, Supplement 4*, 23, 67-77.
- Klieme, E. (1997). *Gender-related differences in mathematical abilities: Effect size, spatial mediation, and item content*. Paper vorgestellt auf der 7th Conference of the European Association for Research in Learning and Instruction (EARLI), Athen, Griechenland.
- Köller, O., Daniels, Z., Schnabel, K. U. & Baumert, J. (2000). Kurswahlen von Mädchen und Knaben im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse. *Zeitschrift für Pädagogische Psychologie*, 14, 26–37.
- Krajewski, K. & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year- longitudinal study. *Journal of Experimental Child Psychology*, 103, 516-531.
- Lachance J. & Mazzocco M. (2006). Longitudinal analysis of sex differences in math and special skills in primary school age children. *Learning and Individual Differences*, 16, 195-216.

- Le Corre, M., Van de Walle, G. A., Brannon, E. & Carey, S. (2006). Re-visiting the performance/competence debate in the acquisition of counting as a representation of the positive integers. *Cognitive Psychology*, 52, 130–169.
- Lüdtke, O., Köller, O., Marsh, H. W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology*, 30, 263-285.
- Lummis, M. & Stevenson, H. W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology*, 26, 254-263.
- Marsh, H.W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23, 129–149.
- Marsh, H. W. (1987). The big fish little pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280-295.
- Marsh, H.W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, 81, 417-430.
- Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, 82, 646-656.
- Marsh, H.W. (2005). Big-fish-little-pond effect on academic self concept. *Zeitschrift für Pädagogische Psychologie*, 19, 119-127.
- Marsh, H. W., Barnes, J., Cairns, L. & Tidman, M. (1984). The Self Description

- Questionnaire (SDQ): Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology*, 76, 940-956.
- Marsh, H. W., Byrne, B. M. & Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80, 366-380.
- Marsh, H. W. & Craven, R. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, 1, 133-163.
- Marsh, H. W., Craven, R. & Debus, R. (1991). Self-concepts of young children aged 5 to 8: Their measurement and multidimensional structure. *Journal of Educational Psychology*, 83, 377-392.
- Mead, G. H. (1934). *Mind, self, and society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55, 19–27.
- Möller, J., Pohlmann, B., Köller, O. & Marsh, H.W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79, 1129-1167.
- Möller, J. & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.).

Pädagogische Psychologie (S. 179- 204). Berlin: Springer.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L. & Smith, T. A. (1997). *Mathematics achievement in the primary school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill: Boston College.

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of own attainment, and the understanding that difficult tasks require more ability. *Child Development*, 49, 800-814.

O'Mara, A. J., Marsh, H. W., Craven, R. G. & Debus, R. (2006). Do self-concept interventions make a difference? A synergetic blend of construct validation and meta-analysis. *Educational Psychologist*, 41, 181-206.

OECD (2004). *Learning for tomorrow's world – First results from PISA 2003*. Paris: OECD.

OECD (2006). *Evolution of student interest in science and technology studies – Global science forum*. Paris: OECD.

OECD (2008). *Education at a glance 2008 – OECD indicators*. Paris: OECD.

OECD (2009). *PISA 2006 technical report*. Paris: OECD.

OECD (2010). *PISA 2009 results: What students know and can do – Student performance in reading, mathematics and science (Volume I)*. Paris: OECD.
[Online] Verfügbar unter: <http://dx.doi.org/10.1787/9789264091450-en>
[27.08.2012].

Penner, A. M. (2008). Gender differences in extreme mathematical achievement: An

- international perspective on biological and social factors. *American Journal of Sociology*, 114, 138–170.
- Penner, A. M. & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research*, 37, 239–253.
- Piaget, J. (1965). *The child's conception of number*. New York: Norton.
- Poloczek, S., Karst, K., Praetorius, A.-K. & Lipowsky, F. (2011). Generalisten oder Spezialisten? Bereichsspezifität und leistungsbezogene Zusammenhänge des schulischen Selbstkonzepts von Schulanfängern. *Zeitschrift für Pädagogische Psychologie*, 25, 173-183.
- Rauer, W. & Schuck, K.D. (2004). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen (FEESS 1-2). Handanweisung*. Göttingen: Beltz Test GmbH.
- Rheinberg, F. (2006). *Motivation* (6. Aufl.). Stuttgart: Kohlhammer.
- Ricken, G., Fritz, A. & Balzer, L. (2011). Mathematik und Rechnen – Test zur Erfassung von Konzepten im Vorschulalter (MARKO-D) – ein Beispiel für einen niveauiorientierten Ansatz. *Empirische Sonderpädagogik*, 3, 256-271.
- Riley, M.S., Greeno, J.G. & Heller, J.H. (1983): Development of children's problem-solving ability in arithmetic. In Ginsburg, H.P. (Hrsg.): *The Development of mathematical thinking* (S. 153-196). New York: Academic Press.
- Roffman, J. L., Marci, C. D., Glick, D. M., Dougherty, D. D. & Rauch, S. L. (2005).

Neuroimaging and the functional neuroanatomy of psychotherapy.

Psychological Medicine, 35, 1385-1398.

Rost, D.H., Sparfeldt, J.R. & Schilling, S.R. (2007). *Differentielles Schulisches Selbstkonzept-Gitter mit Skala zur Erfassung des Selbstkonzepts schulischer Leistungen und Fähigkeiten (DISK-Gitter mit SKSLF-8)*. Manual. Göttingen: Hogrefe.

Schilling, S., Sparfeldt, J. R. & Rost, D. (2006). Facetten schulischen Selbstkonzepts. Welchen Unterschied macht das Geschlecht? *Zeitschrift für Pädagogische Psychologie*, 20, 9-18.

Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2002). *Skalen zur Erfassung des schulischen Selbstkonzepts (SESSKO)*. Manual. Göttingen: Hogrefe.

Schuchardt, K. & Mähler, C. (2010). Unterscheiden sich Subgruppen rechengestörter Kinder in ihrer Arbeitsgedächtniskapazität, im basalen arithmetischen Faktenwissen und in den numerischen Basiskompetenzen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42, 217–225.

Schwippert, K., Bos, W., Lankes, E.-M. (2003): *Heterogenität und Chancengleichheit am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. In: Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G., & Valtin, R. (Eds.): *Erste Ergebnisse aus IGLU: Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.

- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: validation of construct interpretation. *Review of Educational Research*, 46, 407-441.
- Sherman, J. & Fennema, E. (1977). The study of mathematics by high school girls and boys: Related variables. *American Educational Research Journal*, 14, 159-168.
- Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science: A critical review. *American Psychologist*, 60, 950-958.
- Spelke, E. S., Kestenbaum, R., Simons, D. & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *The British Journal of Developmental Psychology*, 13, 113-142.
- Stanat, P. & Kunter, M. (2001). Geschlechterunterschiede in Basiskompetenzen. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann, & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 249-269). Opladen: Leske + Budrich.
- Steffe, L. P. (1992). Learning stages in the construction of the number sequence. In J. Bideaud, C. Meljac, & J. Fischer (Hrsg.), *Pathways to number: Children's developing numerical abilities*, (S. 83-98). Hillsdale, NJ: Lawrence Erlbaum.
- Steinmayr, R. & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, 22, 185-209.
- Stern, E. (1992): Spontaneous use of conceptual mathematical knowledge in elementary school children. *Contemporary Educational Psychology*, 17,

266–277.

Stipek, D. J. & Hoffmann, J. M. (1980). Development of children's performance-related judgments. *Child Development*, 51, 912-914.

Struwe, U. (2005). *idee_it Begleitforschung: Frauen und Männer in IT-Ausbildung und –Beruf*. Bielefeld: Frauen geben Technik neue Impulse e.V. [Online]
Verfügbar unter: <http://www.idee-it.de/var/storage/original/application/phpxPfTrB.pdf> [27.08.2012]

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92, 144-151.

Tiedemann, J. & Faber, G. (1994). Mädchen und Grundschulmathematik: Ergebnisse einer Längsschnittuntersuchung zu ausgewählten geschlechtsbezogenen Unterschieden in der Leistungsentwicklung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 26, 101-111.

Tiedemann, J. & Faber, G. (1995). Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattributionen im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 61-71.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O. & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth grade mathematics. *Journal of Educational Psychology*, 98, 788-806.

Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The relations between self-

- beliefs and academic achievement: A systematic review. *Educational Psychologist*, 39, 111-13.
- Valeski, T. N. & Stipek, D. J. (2001). Young children's feelings about school. *Child development*, 72, 1198-1213.
- Valtin, R., Wagner, C. & Schwippert, K. (2005). Schülerinnen und Schüler am Ende der vierten Klasse - schulische Leistungen, lernbezogene Einstellungen und außerschulische Lernbedingungen. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU - Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 187-238). Münster: Waxmann.
- Van den Heuvel-Panhuizen, M. (2004). Girls' and boys' problems: Gender differences in solving problems in primary school mathematics in the Netherlands. In B. Clarke, D. M. Clarke, G. Emanuelsson, B. Johansson, D. V. Lambdin, F. K. Lester, A. Wallby & K. Wallby (Hrsg.), *International perspectives on learning and teaching mathematics* (S. 237-252). Göteborg: National Center for Mathematics Education.
- Van Houtte, M. (2004). Why boys achieve less at school than girls: The difference between boys' and girls' academic culture. *Educational Studies*, 30, 159-173.
- Van Oeffelen, M.P. & Vos, P.G. (1982) A probabilistic model for the discrimination of visual number. *Perception and Psychophysics*, 32, 163-170.
- Wagner, C. & Valtin R. (2003). Noten oder Verbalbeurteilung? Die Wirkung unterschiedlicher Bewertungsformen auf die schulische Entwicklung von

- Grundschulkindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 35, 27-36.
- Walther, G., Geiser, H., Langeheine, R. & Lobemeier, K. (2003). Mathematische Kompetenzen am Ende der vierten Jahrgangsstufe. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU* (S. 189-226). Münster: Waxmann.
- Weinhold Zulauf, M., Schweiter, M. & von Aster, M. (2003). Das Kindergartenalter: Sensitive Periode für die Entwicklung numerischer Fertigkeiten. *Kindheit und Entwicklung*, 12, 222–230.
- Wigfield, A., Eccles, J.S., Yoon, K., Harold, R.D., Arbretton, A., Freedman-Doan, C. & Blumenfeld, P.C. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89, 451-469.
- Wigfield, A. & Karpathian M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, 26, 233-261.
- Winkelmann, H., van den Heuvel-Panhuizen, M. & Robitzsch, A. (2008). Gender differences in the mathematics achievements of German primary school students: results from a German large-scale study. *ZDM Mathematics Education*, 40, 601-616.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749-750.
- Wynn, K. (1995). Infants possess a system of numerical knowledge. *Current Directions in Psychological Science*, 4, 172-177.

Xu, F. & Arriaga, R. I. (2007) Number discrimination in 10-month-old infants.

British Journal of Developmental Psychology, 25, 103-108.

Xu, F. & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants.

Cognition, 74, B1-B11.

Zimmer, K., Burba, D. & Rost, J. (2004). Kompetenzen von Jungen und Mädchen.

In Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. & Schiefele, U. (Hrsg.). PISA 2003.

Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs (S. 211-223). Münster: Waxmann.

Zohar, A. & Gershikov, A. (2008). Gender and performance in mathematical tasks:

Does the context make a difference? *International Journal of Science and Mathematics Education*, 6, 677-693.

A Selbstkonzepterfassung

A.1 Instruktion

Teil A:

„Ich möchte gerne wissen, wie du darüber denkst, was du so in der Schule machst. Deshalb werde ich dir gleich ein paar Sätze vorlesen und du sagst mir, wie sehr das für dich stimmt. Wir probieren es mal: Ich kann gut malen. Stimmt das ganz genau, stimmt das oft, stimmt das manchmal, stimmt das selten oder stimmt das gar nicht?“

Während die Antwortmöglichkeiten vorgelesen werden, bitte auf das jeweilige Smiley zeigen.

Falls das Kind nicht sofort auf ein Smiley zeigt, bitte die Frage wiederholen. Falls ein Kind die Frage nicht versteht, kann man es auch anders erklären, z.B. zeichnen anstatt malen.

Wenn das Kind auf ein Smiley gezeigt hat, bitte noch einmal nachfragen, z.B. **„Also du kannst manchmal gut malen?“** und auf die Bestätigung des Kindes warten.

Der gleiche Ablauf erfolgt im 2. Teil, in dem sich die Antwortmöglichkeiten ändern. **„Sag mal, wie gut bist du, verglichen mit den anderen Kindern in deiner Klasse im Malen?“**

Bist du der Beste, bei den Besten, bist du mittelgut, bei den Schlechteren oder der Schlechteste?“

Während die Antwortmöglichkeiten vorgelesen werden, bitte auf das jeweilige Smiley zeigen.

Nach Antwort des Kindes, wieder nachfragen zur Absicherung des Verständnisses.

Teil B:

Zum Vergleich der Kinder liegen immer zwei Namenskarten auf dem Tisch. Das zu testende Kind und jeweils das Kind $X+2$; $X+1$ etc.

Es wird jedes Mal die Frage gestellt: **„Was denkst du, wer ist besser im Rechnen; du oder $X+2$...?“**

Dies wird mit dem Lesen und Schreiben fortgeführt.

Es gibt keine Gleichheit, d.h. das Kind muss sich für einen Namen entscheiden.

Verabschiedung

„Vielen Dank, dass du mir alles erzählt hast. Ich fand es sehr interessant.“

A.2 Smileys zur Selbstkonzepterfassung

